



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 128108

To: Sarvamangala Devi

Location: REM 3C18

Art Unit: 1645

Wednesday, August 04, 2004

Case Serial Number: 09/700712

From: Beverly Shears

Location: Remsen Bldg.

RM 1A54

Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes

Shears, Beverly

128108

From: Devi, Sarvamangala
Sent: Monday, July 26, 2004 12:29 PM
To: Shears, Beverly
Subject: 09/700,712

Beverly:

Would you please perform a search for SEQ ID NO: 1 in commercial sequence and interference databases in application SN 09/700,712?

Please include a search for the inventors: Nils Carlin and Michael R. Lebens; and a text search for: A thy A or thymine A strain of Vibrio cholerae comprising a site directed mutagenesis in the thy A gene.

Thanks.

S. DEVI, Ph.D.
AU 1645
Rems - 3C18



STAFF USE ONLY

Date completed: 08-02-04

Searcher: Beverly 2528

Terminal time: 20 / 11

Elapsed time: _____

CPU time: _____

Total time: 23 / 21

Number of Searches: _____

Number of Databases: 3

Search Site

____ STIC

____ CM-1

____ Pre-S

Type of Search

____ N.A. Sequence

____ A.A. Sequence

____ Structure

____ Bibliographic

Vendors

____ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

____ DARC/Questel

____ Other CGA

This Page Blank (uspis,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 12:17:10 ; Search time 4825 Seconds
(without alignments)
18003.955 Million cell updates/sec

Title: US-09-700-712A-1

Perfect score: 2909 1 gagaaggttgcgtcgcgc.....cagcaaatgcgcgtcgag 2909

Sequence: 1 gagaaggttgcgtcgcgc.....cagcaaatgcgcgtcgag 2909

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbo:*
5: em_estbp:*
6: em_estbr:*
7: em_estbs:*
8: em_estbt:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estgm:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gsa1:*
29: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	249.8	8.6	1128	BZ577690 msh2_5528
2	194.8	6.7	805	BZ556310 pacs1-60
3	191.8	6.6	853	BZ578932 msh2_6070
4	161	5.5	1141	BZ559267 pacs2-164

Result No.	Score	Query Match Length	ID	Description
5	152	5.2	1375	BZ553447 pacs1-60
6	115.6	4.0	406	BZ561227 pscu
7	111.2	3.8	1034	BZ561697 pacs2-164
8	111	3.8	878	CD379073 PTM004242
9	111	3.8	921	CD382942 PTM008112
10	106.2	3.7	855	CD376447 PTM01614
11	104	3.6	843	CA588286 hab94911
12	103.8	3.6	813	CA588344 hab50601
13	103.2	3.5	828	CA475083 AGENCOURT
14	96	3.3	685	BW039803 BW039803
15	96	3.3	686	BW039194 BW039194
16	96	3.3	801	CA588156 hab39207
17	92.4	3.2	662	CA589861 hab39207
18	89	3.1	522	BW061136 BW061136
19	89	3.1	632	BW295886 BW295886
20	88.8	3.1	524	CD336403 STRPu537
21	86.6	3.0	591	BG732886 346545 MA
22	85	2.9	683	AW423104 f168d06.X
23	84.8	2.9	700	CA353127 624588 NC
24	84.8	2.9	712	CA373071 647095 NC
25	84.6	2.9	483	CG637075 OST362450
26	84.6	2.9	606	AW546106 L0003F03
27	84.6	2.9	668	BY746793 BY746793
28	84.6	2.9	976	BY746114 BY746114
29	84.6	2.9	2255	AK088606 Mus muscu
30	84.6	2.9	2256	AK088134 Mus muscu
31	84.6	2.9	3128	AK035663 Mus muscu
32	84	2.9	677	CB420463 593420 MA
33	83.8	2.9	756	B1078307 602872658
34	82.8	2.8	776	CF995901 AGENCOURT
35	82.4	2.8	528	CD330233 STRPu537
36	82	2.8	748	B1055160 UI-M-FD0
37	81.8	2.8	673	CD313851 STRPu537
38	81.4	2.8	407	CD330290 STRPu537
39	81.4	2.8	1121	AK011435 Mus muscu
40	80.8	2.8	945	BF783860 602109784
41	80.8	2.8	2486	AK004832 Mus muscu
42	80.2	2.8	952	BF782519 602106535
43	80.2	2.8	1024	BM924499 AGENCOURT
44	79.8	2.7	974	BF607905 MY1.00084
45	79.2	2.7	567	AW108416 um30604.X

ALIGNMENTS

RESULT 1
BZ577690 1128 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_5528.y2 msh Pseudomonas aeruginosa genomic clone msh2_5528,
DEFINITION genomic survey sequence.
ACCESSION BZ577690
VERSION BZ577690.1 GI:27212751
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 1128)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,U.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence Variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) in press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers

source
1..1128
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msb2_5528"
/clone.lib="msb"
/note="Environmental isolate. Whole genomic shotgun library."

Query Match 8.6%; Score 249.8; DB 28; Length 1128;
Best Local Similarity 59.7%; Pred. No. 1.9e-61;
Matches 453; Conservative 0; Mismatches 301; Indels 5; Gaps 2;

29 CTGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 88
29 CTGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 88
93 CTGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 152
93 CTGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 152
89 TATGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 148
153 TATGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 212
153 TATGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 212
149 GATGCGCGGCGAGTGTGAGCGCGGAGTGTGAGCGCGGAGTGTGAGCGGCTT 208
213 AAGCGCTTGCACCGGAGCTGAGCGGAGTGTGAGCGCGGAGTGTGAGCGGCTT 212
209 TTAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
273 TCGGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 332
269 CTGAGTTTCCCATATTTGACCCCGTATGTTTTCATGAGCCCTCAGGCGGTG 328
333 ATGCGCAACCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 392
329 TTAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 388
393 CTGCGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 452
389 GATGCGCGGAGTGTGAGCGCGGAGTGTGAGCGCGGAGTGTGAGCGGAGTGTG 448
453 CAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 512
449 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 508
513 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
509 CTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
573 CCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 632
569 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
633 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 692
629 TCCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
633 TC--GGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
669 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 745
751 CCGGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 810
746 TTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
811 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 849

RESULT 2 805 bp DNA linear GSS 17-DEC-2002
B2556310
LOCUS pacsl-60.5617.x1 pacsl-60 Pseudomonas aeruginosa genomic clone
DEFINITION pacsl-60.5617, genomic survey sequence.

ACCESSION B2556310 GI:27166959
VERSION B2556310.1
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 805)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1..805
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60.5617"
/clone.lib="pacsl-60"
/note="Clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN
Query Match 6.7%; Score 194.8; DB 28; Length 805;
Best Local Similarity 60.6%; Pred. No. 2.5e-45;
Matches 319; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

47 GACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 106
103 GATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 162
107 GTGCGTTTCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 166
163 ATCGGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 222
167 TGAAGCGGAGCAAGTCTGACTGTTTATTCGCGGCTTTTAAAGTGTGATGCGGT 226
223 TGAAGCGGAGCAAGTCTGACTGTTTATTCGCGGCTTTTAAAGTGTGATGCGGT 282
227 GCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 286
283 GCGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 342
287 TTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 346
343 ATCTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 402
347 GCGATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 406
403 GCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 462
407 GCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 466
463 GCGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 522
467 CTTTGGGAGCGATAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
523 CTGTGGGAGCGATAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 582
527 CCGGCGCATCTTCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
583 CCGGCGCATCTTCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628

```

RESULT 3
LOCUS      BZ578932          853 bp    DNA          linear    GSS 17-DEC-2002
DEFINITION msh2_6070.y2 msh Pseudomonas aeruginosa genomic clone msh2_6070,
            genomic survey sequence.
ACCESSION  BZ578932
VERSION    BZ578932.1  GI:27213993
KEYWORDS   GSS.
SOURCE     Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE  1 (bases 1 to 853)
            Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
            Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
            J. Bacteriol. (2002) in press
            Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: Shotgun.
FEATURES
            source
            1..853
               location/Qualifiers
               /organism="Pseudomonas aeruginosa"
               /mol_type="genomic DNA"
               /strain="M5H"
               /db_xref="taxon:287"
               /clone="msh2_6070"
               /clone_11b="msh"
               /note="Environmental isolate. Whole genomic shotgun
               library."
ORIGIN
Query Match      6.6%; Score 191.8; DB 28; Length 853;
Best Local Similarity 60.0%; Pred. No. 2e-44;
Matches 336; Conservative 0; Mismatches 223; Indels 1; Gaps 1;
QY      29  CTGACGTTTCCCATATATGACCCCGTATGTTTTCATGCGCCCTCTAGCGGTGCGCTGG 88
DB      168  CTGACGATCCCGACATCGATCCGTTGCGCTGCGCCATCGCCCGCTGAAGATCACTGG 227
QY      89  TATGGCTTGATGATTTGTTGGGTTTCTTTTGTATGTTGTTGCAATCGCCGAGCG 148
DB      228  TAGGGGCTGATGATCGATCGGGATCGGGGCGCCCTGCTGCGCTGCGCCCGGATG 287
QY      149  GATCGCGCGGCGAGTGTGGACGCGTGAAGAGTCTTCTGACTTGTATTGCGCGGCTTT 208
DB      288  AAGCGCTTCAACCGACCTGACCAAGAGAGCGCTTCCACCTGCTCTGCGTGGCG 347
QY      209  TTAGTGTAGTGAATCGGTGGCGAGTTGTTATGATCTTCACAATTTTGAATCTGTC 268
DB      348  TGGGGCTGATCTCTGGTGGCGCGCTGAGCTGATGCTGTTCAACCTGGAGCATGAC 407
QY      269  CTGCTGACCCCTTTATTTATTAAGATGTGAGTGGCGGCAATGCTCTTCCACGCGCG 328
DB      408  ATGCGCAACCGGACGCTGATCTTGAAGTCTGGAAGGGCGGCAATGCTTCCATGCGCG 467
QY      329  TTATTTGGTGTGATCAACCGGCAATGTTCTGTTATGCGCGTAAACCAAGCAAGCTTCTT 388
DB      468  CTGCTGCGCTGATGCTGCGCGTCTGATGTTGCGCAAGCGCCATGCGCAAGCTTCTTC 527
QY      389  GGTGTGGCCGATTTTGTGCGCTTTAGTGCATTCGTTTGGGATGGGACGATACGAT 448
DB      528  CAGCTGATGAGCTTCAATCGCCCGCTGTTGCCATCGGCTGGGCGCGCGGCAATCGCG 587
QY      449  AACTTATGATAGTGAACCTTTGGGACGATGACGATGTGCTTGGGCTTTTGTATTC 508

```

```

DB      588  AACTTATGATAGTGAACCTTTGGGACGATGACGATGTGCTTGGGCGCATGCTTTC 647
QY      509  CCTATGTTGGCCCACTGCGCGGCAATCCTTCAACGTTTATGAATTCGCCCTTAGAAGGC 568
DB      648  CCCAAACG-CGCCCGCTGCGCGGCAATCCTTCAACGTTTATGAATTCGCCCTTAGAAGGC 706
QY      569  GTGCTTCTGCTTATTTCT 588
DB      707  CGTGGCTGTGCTATCTCT 726

RESULT 4
LOCUS      BZ559267          1141 bp    DNA          linear    GSS 17-DEC-2002
DEFINITION pacs2-164_1355.s1 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_1355, genomic survey sequence.
ACCESSION  BZ559267
VERSION    BZ559267.1  GI:27175318
KEYWORDS   GSS.
SOURCE     Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE  1 (bases 1 to 1141)
            Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
            Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
            J. Bacteriol. (2002) in press
            Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: Shotgun.
FEATURES
            source
            1..1141
               location/Qualifiers
               /organism="Pseudomonas aeruginosa"
               /mol_type="genomic DNA"
               /strain="2-164"
               /db_xref="taxon:287"
               /clone="pacs2-164_1355"
               /clone_11b="pacs2-164"
               /note="Clinical isolate 2-164 Whole genomic shotgun
               library."
ORIGIN
Query Match      5.5%; Score 161; DB 28; Length 1141;
Best Local Similarity 64.6%; Pred. No. 2.7e-35;
Matches 272; Conservative 0; Mismatches 145; Indels 4; Gaps 2;
QY      380  ACCCTCTTGGTGGCGGCAATTTTGTGCGCCCTTATGAGCATTCGGTTGGGATGGA 439
DB      628  AGCTTCTTCAAGCTGATGATTTATCGCCCGCTGTTGCCATCGGCTGGGCGCGGG 569
QY      440  CGTATCGTAACTTTATGATGTAAGTCTTTGGGAGCAGTAAAGATGTCCTTGGGCT 499
DB      568  CGCATCGGCACTTCACTCACTCGGAAGTGGGGCAAGTCAAGATGTCCTTGGGCG 509
QY      500  TTTGATTCCTTATGATGTTGGCGGCACTGCGCGGCAATCTTCAAGCTTATGATTTGCC 559
DB      508  ATGCTTTCCTCAACGCGGGGAC-GTGGCGGGGATCTCTGCACTGTACAGATTGGCC 450
QY      560  TTAGAAGCGTGTCTGTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 619
DB      449  CTGGAAGCGTGGGGGCTGTTCTATCTCTGCTGTTTCAACCGGCAAGCGGCGGAC 390
QY      620  GGCAGCGTATCGGACGTTTATTTAGCTGATACGTTACATTTCCGCTTCTTGGAAATAC 679
DB      389  GCTTGGCTTCGCGGCTGTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 330

```

Qy	680	GTCCGAGACAGATGCTCAGTTGG---TCGTGTGGCTTCATTTCAGGGGCA	736
Db	329	GTCCGGGTGCGGATATCCAGCTGCGTACTCGCTTGGGCTTGGACCAATGGGCG	270
Qy	737	ATCCTCCCTTACCTATGTGTGATCATCGGTATTTTGATGATGTTTGGTCTTAAAGCG	796
Db	269	GTGCTCTGGTACCGATGTGTCTGGCCGCGCATTGCGCTGATGTGTCTGGGCTTAAAGCGCG	210
Qy	797	G 797	
Db	209	G 209	

RESULT 5	BZ553447	1375 bp	DNA	linear	GSS 17-DEC-2002
LOCUS	BZ553447				
DEFINITION	pacsl1-60_4149.x1				Pseudomonas aeruginosa genomic clone
	pacsl1-60_4149,				genomic survey sequence.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Bacteriia: Protobactetia: Gemmaproteobactetia: Pseudomonadates: Pseudomonadaceae: Pseudomonas.
1 (bases 1 to 1375)
Spencer, D. H., Raymond, C. K., Smith, E. E., Sims, E. E., Hastings, M.,
Burns, J. O., Kaul, R. and Olsen, M. V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) in press
Contact: Chris K. Raymond

FEATURES	Location/Qualifiers
source	1..1375
	/organism="Pseudomonas aeruginosa"
	/mol_type="genomic DNA"
	/strain="1-60"
	/db_xref="taxon:287"
	/clone="pacs1-60 4149"
	/clone_1ib="pacs1-60"
	/note="clinical isolate 1-60 Whole genomic shotgun library."

Db 324 CCAACGGGGCGCCGCTCCGGGGATCCCTGAAAGTTGACCAAGTTGGCCCTGAAAAGCC 383
 Qy 570 TGGTTCGTGTTCTTAATCTTAATGAGTTTATGTAACCTCGGCGCTAGGCACGGAT 629
 Db 384 TGGCAGCTGTTGCTATCCTCTGCGTGGCCACCCGATGCCAGGATCCACGCGCTTTGCT 443
 Qy 630 CCGACACTGTTTATGCTGGAATACGGATCATTCGCTTCTTGTAATACGTCCTGAGC 689
 Db 444 CTGGCCGTGTTGCTGCTGATACCGGATCTACCGCTTTGATGGCGAATTCACCGCTGGC 503
 Qy 690 CAGATGCTCAATGG---GTCGTGTTGAGTGGCTCATTTAATAGGGGAAATCTCTGCT 745
 Db 504 CGCATGCCCACTGTCTACTCTGCACGTAGGCTTGCTGACCATGGGCCAAGAGCTGCAG 563
 Qy 747 TACCTATGAGTATCATCGTATTT 770
 Db 564 TACCCGTGAATGACCCGTTCCT 587

Query Match	4.0%	Score 115.6	DB 29	Length 406
Best Local Similarity	59.8%	Pred. No. 3.8e-22		
Matches 234	Conservative	0	Mismatches 145	Indels 12
			Gaps	2
QY	399	ATTGTTGTCCTCCCTTTAGTGCATTCGGTTTGCGGATGCGACGATTCGGTAATTATGA	458	
Db	15	ATTTCCTTCCCGCGCTGGTGGCGGATGGCGNTGGGATCCCGGCGTATCGGCAACTTCATCA	74	
QY	459	ATATGTAACCTTTGGGAGACAGTAAACGATGTGCTTTGGCTTTTGATTTCCC-----	510	
Db	75	AGCGCGAGCTGGGGGCAACCCAGCAGTGGCAAGGGCGATGATCTTTCCGCCGCTTCA	134	
QY	511	-TATATGTGGCCCACTGCCGCGCATTCCTTCAAGTTATGAAATTCGCTTAGAAGGCG	569	

Db 135 GCGACCCAGCCAGCTCCACGACCCGTCGACGCTGACAGTTGCGCCTGGAGGGCG 194
 QY 570 TGGTGTCTTATCTTAATTGTTATTTGTAACCTGTCGCTGAGGAGGAT 629
 Db 195 TGGACATGTTGTCATCTCTGCTGTTCTGCGCAGCCGCGCCAGCATTGCGGAT 254
 QY 630 CCGGACTGTTTATAGTGAATACGATACATTCGCTTCCTTGAATACGTCGAGC 689
 Db 255 CGGACATGTTGCGCTGTTCTACGCGCATCTTCGCTTCATGTCGAATTGTCGCGGTGC 314
 QY 690 CAGATGCTGATGGG---TCTGTTGGTGGCTTCATTCAATGGGCAATCTCTCT 746
 Db 315 CGATGCGCAGCTGGGCTATATGCGCTGGGCTGGCTGACCATGCGCATTCGTGTGC 374
 QY 747 TACCTATGATGATCATCGATATTTTATGAT 777
 Db 375 TGCCGATGATCTGCTGGCCTTGCCCTGAT 405

RESULT 7
 B2561697 1034 bp DNA linear GSS 17-DEC-2002
 LOCUS pacs2-164_3389.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
 DEFINITION pacs2-164_3389, genomic survey sequence.
 ACCESSION B2561697
 VERSION B2561697.1 GI:27182207
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1034)
 Spencer, D.H., Raymond, C.K., Smith, B.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
 Whole-Genome-Sequence Variation among Multiple Isolates of Pseudomonas aeruginosa Library J. Bacteriol. (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers
 source 1..1034
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="2-164"
 /db_xref="taxon:287"
 /clone="pacs2-164_3389"
 /clone_1id="pacs2-164"
 /note="clinical isolate 2-164 whole genomic shotgun library."

ORIGIN
 Query Match 3.8%; Score 111.2; DB 28; Length 1034;
 Best Local Similarity 53.9%; Pred. No. 1.2e-20;
 Matches 342; Conservative 0; Mismatches 278; Indels 14; Gaps 5;

QY 206 TTTTAGGTAGTATGATGATGAGGAGGAGTGGTATGATCTTCTACATTTTATGATCG 265
 Db 550 CTGGGGGGGTATATCTCCGATGCGCGCTGGGCTACAGTGTGTTCTACATCGGACAG 609
 QY 266 TTCCTGCTGACCCCTCTTATTTATTTAAAGTGTGAGATGCGGCGATCTCTCCAGCC 325
 Db 610 TACATGCGCAACCCACGCTGATCTTCAGAGTCTGGGAAGCGGCGATCTCTTCATGC 669
 QY 326 GCGTATTTGGGTGATGATGACCGGCATGTTCTGTATGATGAGCGGTAACCAAGCAGCTTC 385
 Db 670 GGGCTGCTCGGGTATGCTGCGGTCTGCTTGGTTTGGCAGAGGCGCATGGCAAGAGC 729
 QY 386 TTGGTGTGGCGCA---TTTGTGGCCCTTTAGTGCATTCGCTTTGGGGATGGAC 440
 Db 730 TTCCTTCAGCCTGAGGAGACTTCATTCGCGCCGTTGGTCCCATGGGCTGGGCGCGCGC 789
 QY 441 GTATGCGTAACTTTATGAAATGAACTTTGGGACGAG---TAACGATGTCCTTGGG 497
 Db 790 GCATTGGGAATCTTTATTAATCTGAACTTGAGGGAAGTCAATGATGAGCCCGGCGC 849
 QY 498 CTTTGTATTCCTTAATGATGATGAGCCGACATGCGCGCATCTTCAAGCTTTATGAA-TTC 556
 Db 850 CCAAGGCTTTCCCAAGCTCGGCCCTTTCCGGGGCATCCCTTACAAATTGCCAATTTC 909
 QY 557 GCCTTAGAAGGCGGTGCTCTGTTCTTTA--TTCTTAATGTTTATTTGTAACCTGTC 614
 Db 910 GCCTTAGAAGGCGGGGCGCGCTTTTAAATCTTGGGTGTTAACCACCAAGCCCGGCG 969
 QY 615 CGTAGGACGAGCTATCCGAGCTGTTTATGCTGG 648
 Db 970 CGCCCGCCTTGCTTCGCTCTTTCCGCGCTTG 1003

RESULT 8
 CD379073 878 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM004242 Pheodactylum tricornutum Uni-Zap XR Pheodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD379073
 VERSION CD379073.1 GI:31254687
 KEYWORDS EST.
 SOURCE Pheodactylum tricornutum
 ORGANISM Pheodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae; Bacillariophycidae; Naviculales; Pheodactylaceae; Pheodactylum.

REFERENCE 1 (bases 1 to 878)
 Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 Genome properties of the diatom Pheodactylum tricornutum
 Plant Physiol. 129 (3), 993-1002 (2002)
 22111123
 MEDLINE 12114555
 PUBMED 12114555
 CONTACT: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://avesthagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source 1..878
 Location/Qualifiers
 /organism="Pheodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /db_xref="taxon:2850"
 /clone_1id="Pheodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN

Query Match 3.8%; Score 111; DB 14; Length 878;
 Best Local Similarity 59.8%; Pred. No. 1.2e-20;
 Matches 186; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

2474 ATGATTTCTAGAGGTAGAGATCAGACCGGAGAAAGATTGTAAGAAATTCGTGATGTTG 2533
 332 ATGACTTCTACAGAGAAAGATGATGGCCACCGAAGAAATTAACATATGTTGAGAGGTG 273
 2534 GCACCTGGCGAAGAGACCGGCGAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAG 2593
 272 GCACGAGGAAAGACCGCTGAGCTGATCCCATCGCCATTGGCCATAGAGAGATT 213
 2594 GTATTGGTCAAGTAGTACCAATATTGGACCCATCAACATAGCAATCGGGTTCAACC 2653
 212 GTGTGGTGAAGTGAAGTACCAATATTCGCTCCCATGATCATGTAATTCCTTGGCGGAG 153
 2554 GGTAAACCAACCGGCAAGACGACCAATTAATAGAACTACCGTCTTGAAGATTGAATG 2713
 152 GAGACAGCGGATCGACCAAGACCAACATCGAGTCTTTGAGAGAAAGCTGAGC 93
 2714 AGTCCGCTTGCACCTAACAATCATCATCTGCAATGGTGGGAAAGCAAAATTCAAAT 2773
 92 AATACAGTCGAAAGATACCAACCATTAAGACGACATGGGATTCGTGTCGCCGAGC 33
 2774 AGAATTGGC 2784
 32 AACTTCGGC 22

RESULT 9
 CD382942/c 921 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM08112 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD382942
 VERSION CD382942.1 GI:31258556
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 921)
 REFERENCE Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://avesthagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source location/Qualifiers
 1..921
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

ORIGIN
 Query Match 3.8%; Score 111; DB 14; Length 921;
 Best Local Similarity 59.8%; Pred. No. 1.3e-20;
 Matches 186; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

2474 ATCATTTCTAGAGGTAGAGATCAGACCGGAGAAAGATTGAAATAATCGTGAATGTTG 2533
 323 ATGACTTCTACAGAGAAAGATGATGGCCACCTGACAGAAATTAACATATTCGTGAGCGTTG 264
 2534 GCACCTGGCGAAGACCGGCGAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAG 2593

DB 353 ATGACTTCTACAGAGAAAGATGATGGCCACCGAGAAATTAACATATGCTGACCGTTG 294
 2534 GCACCTGGCGAAGACCGGCGAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAG 2593
 293 GCACGAGGAAAGACCGCTGAGCTGATCCCATCGCCATTGGCCCATAGAGACGATT 234
 2594 GTATTGGTCAAGTAGTACCAATATTGGACCCATCAACATAGCAATCGGGTTTCAACC 2653
 223 GTGTGGTGAAGTGAAGTACCAATATTCGCTCCCATGATCATGTAATTCCTTGGCGGAG 174
 2654 GGTAAACCAACCGGCAAGACGACCAATTAATAGAACTACCGGCTTGAAGATTGAATC 2713
 173 GAGACAGCGGATCGACCAAGACCAACATTCAGAGTCTTTGAGAGAAAGACTGAGC 114
 2714 AGTCCGCTTGCACCTAACAATCATCATCTGCAATGGTGGGAAAGCAAAATTCAAAT 2773
 113 AATACAGTCGAAAGATACCAACCATTAAGACGACATGGGATTCGTGTCGCCGAGC 54
 2774 AGAATTGGC 2784
 53 AACTTCGGC 43

RESULT 10
 CD376447/c 855 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM01614 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD376447
 VERSION CD376447.1 GI:31252061
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 855)
 REFERENCE Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://avesthagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source location/Qualifiers
 1..855
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

ORIGIN
 Query Match 3.7%; Score 106.2; DB 14; Length 855;
 Best Local Similarity 58.8%; Pred. No. 3.1e-19;
 Matches 183; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

2474 ATCATTTCTAGAGGTAGAGATCAGACCGGAGAAAGATTGAAATAATCGTGAATGTTG 2533
 323 ATGACTTCTACAGAGAAAGATGATGGCCACCTGACAGAAATTAACATATTCGTGAGCGTTG 264
 2534 GCACCTGGCGAAGACCGGCGAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAG 2593

Db 263 GCACCGAGGAAAGACGCTGAGTTGATCCCATCCGCCATTGGCCCATAGGAGAGATG 204

QY 2594 GTATTGGTCAAGTAGTACCAATATTGGACCCCATCCATAGGAATCGCGTTTCAACC 2653

Db 203 GTGTGGTGAAGTGAAGTACCAATATTGGTCCCATGATCATGTAAATACCTTGGCGAGC 144

QY 2654 GGTAAACCGGAGGAGACCAACCAATATAGAGTCCGCGTGTAGAGATTGAATC 2713

Db 143 GAGCAGCGGATCCGACCAAGACACACATCGAGTGTGTTGAAGCAGCTGAGC 84

QY 2714 AGTCCCTTCGCACTTAACCATCATCATCTGCAATTTGGTGGAGCAATTTGAAT 2773

Db 83 AATACATTCGCAAGAAATACCAACCATTAAGACGAGCGATGGGATTCGTGTGCGCGAAG 24

QY 2774 AGAATTTGGC 2784

Db 23 AACTCTCCGCG 13

RESULT 11 843 bp mRNA linear EST 19-NOV-2002

CA588286/hdb49a11.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone IMAGE:6351812 5' similar to TR:Q91237 Q91237 NM/PI COTRANSPORT SYSTEM PROTEIN. ; mRNA sequence.

ACCESSION CA588286

VERSION CA588286.1 GI:25132864

KEYWORDS EST.

SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM Takifugu rubripes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Takifugu.

1 (bases 1 to 843)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

Washu Zebrafish EST Project 1998

Unpublished (1998)

TITLE JOURNAL

COMMENT Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

Library materials provided by G. Elgar (UK MRC HGMP-RC) Library constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Matabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA Sequencing by: Washington University Genome Sequencing Center Clone through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lnl.gov

Seq primer: T3 ET from Amersham

High quality sequence stop: 496.

FEATURES

Source

1..843

Location/Qualifiers

/organism="Takifugu rubripes"

/mol_type="mRNA"

/db_xref="taxon:31033"

/clone="IMAGE:6351812"

/tissue_type="gut"

/dev_stage="adult"

/lab_host="DB10B (T1 phage-resistant)"

/clone_id="Fugu UT6 adult gut"

/note="Vector: pBluescript-PL; Site 1: PflM I (CCANNNTTGG); Site 2: PflM I (CCANNNTTGG); BamH1-SmaI sites were converted to BamH1-PflM1-SfiI-PflM1 sites (SmaI is destroyed). Other part of the vector is untouched. The

ORIGIN

Query Match 3.6%; Score 104; DB 14; Length 843;

Best Local Similarity 58.9%; Pred. No. 1.4e-18;

Matches 198; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

2427 GCAGAGCGAAGACGAGCGAGTGTAGACTTTCTTAATATGCCAAACATATTTTGAAG 2486

Db 721 GGAAGAGCTAATCAAGAACTGGTGAAGCTGTGAAGAACTCTGTACCACTCCAGAG 662

QY 2487 GTAGAAATCAAGACCGCGAGAGATTGAAAAATCTGTGATGTGCACTGGCGAAG 2546

Db 661 GCAGAAAGATCAACACGACGACGAGTTAAAGATCTGTGACGCTGCCCGAGCAAG 602

QY 2547 CAGGCGGAAACTCTCTTTACAGCGCATATGCGCAGGCTGACGAGATTTGTCACAG 2606

Db 601 CCTGGGAACTCTGTTTGATCCCGCTGATCATATGCGACATGTGTTGGTGAAG 542

QY 2607 TAGTACCAATATTGGACCCCATACCATAGAAATGCGGTTT--CAACCGTAAACGAC 2663

Db 541 AGGTCCGATGTGGCGCCCATGATATCGGACCGCGTGTGACATCATGATTTCCAG 482

QY 2664 CGGCAACGAGACCAACATATAGAAATGACCGTCTGTAGATGTGATGTCGCGTGG 2723

Db 481 AGGACACCACTGTGACCAACAAATAGGAGGAGATCTGAGAGCTTGCACGACAGTGA 422

QY 2724 CCAGTAAACCATCATCATCTGCAATTTGGTGG 2759

Db 421 CTAAACCCCGATCAACGACCCGACGACGAGGTTGG 386

RESULT 12 813 bp mRNA linear EST 19-NOV-2002

CA588344/hdb50c01.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone IMAGE:6351745 5' similar to TR:Q91237 Q91237 NM/PI COTRANSPORT SYSTEM PROTEIN. ; mRNA sequence.

ACCESSION CA588344

VERSION CA588344.1 GI:25132922

KEYWORDS EST.

SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM Takifugu rubripes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Takifugu.

1 (bases 1 to 813)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

Washu Zebrafish EST Project 1998

Unpublished (1998)

TITLE JOURNAL

COMMENT Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

Library materials provided by G. Elgar (UK MRC HGMP-RC) Library constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Matabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA Sequencing by: Washington University Genome Sequencing Center Clone

distribution: Fugu clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.llnl.gov
 Seg primer: T3 RT from Amersham
 High quality sequence stop: 497.
 Location/Qualifiers

FEATURES

1..813

/organism="Takifugu rubripes"
 /mol_type="mRNA"
 /db_xref="taxon:31033"
 /clone="IMAGE:6351745"
 /rname_type="gut"
 /dev_stage="adult"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_id="Fugu UT6 adult gut"
 /note="Vector: pBluescript-FL; Site 1: Pflm I (CCANNNNNTGG); Site 2: Pflm I (CCANNNNNTGG); BamHI-SmaI sites were converted to BamHI-PflmI-SfiI-PflmI sites (SmaI is destroyed). Other part of the vector is untouched. The cDNA is inserted between two PflmI sites in T3 (5') to T7 (3') direction. Library materials provided by G. Elgar (UK MRC HGP-RC) and constructed and donated by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Warabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology)."

ORIGIN

Query Match 3.6%; Score 103.8; DB 14; Length 813;
 Best Local Similarity 59.2%; Pred. No. 1.6e-18;
 Matches 196; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

2432 GGGCAAGACGACCGAGTGGAGCTTCTCTTAATGGCAAAACATCTTCTAGAGTAGG 2491
 713 GACGTATCAAGAACTGGGTGAAGTCTGAAAGAACTCTGTAGCCACCTCCAGAGGAGA 654
 2492 AAGATCGACCGCGAGAAATTGAAATTCGTGATGTGTGCATCGGAGAAAGCACGG 2551
 653 AGATCAACACCGACGACGAGTAAAGAACTGGAAGCGGTGCGCCAGCAAAAGCCCTG 594
 2552 CGAACTCTTTTACAGCGCATATGGCCAAAGCTGACGAGATTTGTTCACAGTAGTA 2611
 593 CGGAACCTGTTTGCATCCCGCTGATCATGTGCCAGATGGTGTGGACAGAGGTC 534
 2612 CCAATATTGGACCCATCATCAGATCGCGGTTT---CAACCGGTAACCCACCGGCA 2668
 533 CGATGTTGGCGCCATATATATGCGACCCCGCTCGACATTCATATTCAGAGGAC 474
 2669 ACGAGACCAACATATATGAAAGTACCGTCTTGAGATTAATCAAGTCCGTCCTACT 2728
 473 ACCATGTGACCAATATGAGAGAGAGTGTGAGCTTGACAGACAGGTGACTAAC 414
 2729 AAACCAATCATCATCCGCAATTGGGTGGG 2759
 413 ACCCGATCACCAAGCCACGAGGTTGG 383

RESULT 13
 CA475083 828 bp mRNA linear EST 12-NOV-2002
 LOCUS CA475083
 DEFINITION AGENCOURT 10670196 NCI CGAP_ZKId1 Danio rerio cDNA clone
 IMAGE:6797422 5', mRNA sequence.
 ACCESSION CA475083
 VERSION CA475083.1 GI:24931435
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 828)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Leonard I. Zon, M.D.
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM14310 row: K column: 21
 High quality sequence stop: 433.
 Location/Qualifiers

FEATURES

1..828

/organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6797422"
 /lab_host="DH10B (T1-resistant)"
 /clone_id="NCI CGAP_ZKId1"
 /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: 0130 dt. Average insert size 1.8 kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI CGAP library."

ORIGIN

Query Match 3.5%; Score 103.2; DB 14; Length 828;
 Best Local Similarity 60.5%; Pred. No. 2.4e-18;
 Matches 187; Conservative 0; Mismatches 119; Indels 3; Gaps 1;

2465 ATGCCAAACATCATTTCTAGAGGTAGGAAGTACGACCGCGAGAGATTGAAAAATCG 2524
 521 ATGCTGTGATGCTTTCCAGCGCGACGACGACGACTGACACCACTTAACATTCG 462
 2525 TGGATGTGGCACTGGCGAAGACGCGGAACTCTTTTACAGCCCATATGCGCAAGG 2584
 461 TGCACTGTGGGACCAAGCAAAAGCTGTTGAACTCTCTCTCTCCCTGCTGATGAGA 402
 2585 CTGACGAGATTTGGTGCACAGTAGTACCATATTTGGACCATCATCATAGATGCGG 2644
 401 GCGACGATGTTGTGTGACAGATGTCCAAATGTTTGAAGCCCATGATGATGGGATGGG 342
 2645 GTT---TGACCGGTAAACCAACCGGCAAGAGCAACATATATGAGTACCGTCTT 2701
 341 GATCCACATCATCAGCAATCCAGAGAAACGACGATGACCAATGAGTAGAGTGGTCTG 282
 2702 GAGCATGATATGTCGCGCTTGCACATTAACCATCATCATCTCTGCAATTGGGTGGAA 2761
 281 GAACCTTGACAGAGACGATACGATTTTCCACACCACTCCCGCACAGGGTTAGAA 222
 2762 GCAATTCGA 2770
 221 AGAACCGCA 213

RESULT 14
 BM039803 685 bp mRNA linear EST 19-OCT-2002
 LOCUS BM039803
 DEFINITION BM039803 Nori Satoh unpublished cDNA library, blood cells Ciona intestinalis cDNA clone c1bd041k19 5', mRNA sequence.
 ACCESSION BM039803
 VERSION BM039803.1 GI:24140498
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cloniidae; Ciona.
 REFERENCE 1 (bases 1 to 685)
 AUTHORS Satoh, Y., Satake, M., Azumi, K., Norioka, M., Shin-i, T., Kohara, Y. and Satoh, N.
 TITLE Expressed genes in Ciona intestinalis (2002)

This Page Blank (uspto)

Mon Aug 2 14:24:10 2004

us-09-700-712a-1.rmpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 15:39:05 / Search time 888 Seconds
(without alignments)
16062.201 Million cell updates/sec

Title: US-09-700-712a-1

Sequence: 1 gagaagtttgatgcctc.....cagcgaatcgcgcag 2909

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:*

1: /cgn2_6/ptodara/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodara/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodara/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodara/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodara/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodara/2/pubpna/PCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodara/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodara/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodara/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodara/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodara/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodara/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodara/2/pubpna/US09C_PUBCOMB.seq:*
14: /cgn2_6/ptodara/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodara/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodara/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodara/2/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodara/2/pubpna/US60_PUBCOMB.seq:*
19: /cgn2_6/ptodara/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112.6	38.2	1149	US-10-282-122A-40937	Sequence 40937, A
2	842.4	29.0	852	US-10-282-122A-40936	Sequence 40936, A
3	744.2	25.6	1830121	US-10-329-960-1	Sequence 1, Appl
4	744.2	25.6	1830121	US-10-329-670-1	Sequence 1, Appl
5	505.4	17.4	852	US-10-282-122A-30692	Sequence 30692, A
6	462.2	15.9	852	US-10-282-122A-22162	Sequence 22162, A
7	403.8	13.9	849	US-10-282-122A-32325	Sequence 32325, A
8	248.2	8.5	876	US-09-741-669-182	Sequence 182, App
9	152.2	5.2	640681	US-09-790-988-1	Sequence 1, Appl
10	142.2	4.9	96109	US-10-672-787-35	Sequence 35, Appl
11	115.4	4.0	2731748	US-10-297-465A-1	Sequence 1, Appl
12	84	2.9	2440	US-10-152-319A-1589	Sequence 1589, Ap
13	77.8	2.7	3950	US-10-152-319A-1588	Sequence 1588, Ap
14	75.4	2.6	2208	US-10-276-774-1331	Sequence 1331, Ap

15	75.4	2.6	2280	US-10-173-999-149	Sequence 149, App
16	75.4	2.6	3952	US-10-116-802-36	Sequence 36, Appl
17	75.4	2.6	4137	US-10-052-664-2	Sequence 2, Appl1
18	73.8	2.5	4111	US-10-257-021-146	Sequence 146, Appl
19	73.8	2.5	4135	US-10-097-340-220	Sequence 220, App
20	72.2	2.5	687	US-10-404-460-258	Sequence 258, App
21	70.2	2.4	1800	US-10-169-395-71	Sequence 71, Appl
22	70.2	2.4	2016	US-10-169-395-81	Sequence 81, Appl
23	70.2	2.4	2124	US-10-108-260A-1011	Sequence 1011, Ap
24	70	2.4	1164	US-09-738-626-5017	Sequence 3017, App
25	70	2.4	1333	US-10-627-476-511	Sequence 511, App
26	70	2.4	3309400	US-09-738-626-511	Sequence 1, Appl1
27	69.2	2.4	969	US-10-282-122A-33224	Sequence 33224, A
28	68.2	2.3	795	US-10-282-122A-42178	Sequence 42178, A
29	67.6	2.3	969	US-10-282-122A-14737	Sequence 14737, A
30	66.4	2.3	792	US-10-282-122A-12730	Sequence 12730, A
31	65.4	2.2	722	US-10-332-859-254	Sequence 254, App
32	64.8	2.2	1478	US-10-084-817-332	Sequence 332, App
33	64.8	2.2	1536	US-09-877-178-11	Sequence 11, Appl
34	64.8	2.2	1536	US-10-342-887-556	Sequence 556, App
35	64.8	2.2	1536	US-10-172-118-556	Sequence 556, App
36	64.8	2.2	1536	US-10-388-560-350	Sequence 350, App
37	64.8	2.2	1536	US-10-159-563-244	Sequence 244, App
38	64.8	2.2	1536	US-10-283-975A-254	Sequence 254, App
39	64.8	2.2	3298	US-10-240-965-200	Sequence 200, App
40	64.6	2.2	795	US-09-741-669-181	Sequence 181, App
41	64.6	2.2	795	US-10-282-122A-6574	Sequence 6574, Ap
42	64.6	2.2	798	US-09-738-626-942	Sequence 942, App
43	64.6	2.2	1200	US-09-954-197-1	Sequence 1, Appl1
44	64.6	2.2	2730	US-09-987-446-1	Sequence 1, Appl1
45	64.4	2.2	969	US-10-282-122A-14234	Sequence 14234, A

ALIGNMENTS

RESULT 1
US-10-282-122A-40937/c
Sequence 40937, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 40937
LENGTH: 1149
TYPE: DNA
ORGANISM: *Vibrio cholerae*
US-10-282-122A-40937

Query Match 38.2%; Score 1112.6; DB 13; Length 1149;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1118; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1783 TTAAGCGGTGAGGCAAGATGCTGCGGGGATGACGCAAAACACACCAATTAAGTACT 1842
DB 1149 TTAAGCGGTGAGGCAAGATGCTGCGGGGATGACGCAAAACACCAATTAAGTACT 1090
QY 1843 CACCACACCAATTTGCTCTTACAGCCCAAGTTAGATGAGCTCAGCACCCTTAATAG 1902
DB 1089 CACCACACCAATTTGCTCTTACAGCCCAAGTTAGATGAGCTCAGCACCCTTAATAG 1030
QY 1903 CAGTTCCGCTAAGAAAGGAATCCGTTAATCAGACCGCTGACCAAGTTAAAGCTTAA 1962
DB 1029 CAGTTCCGCTAAGAAAGGAATCCGTTAATCAGACCGCTGACCAAGTTAAAGCTTAA 970
QY 1963 GTGACACGACCAATTTGCTGAGCAAGCAACGCGAACTCAGCAGACCGGTTGGCG 2022
DB 969 GTGACACGACCAATTTGCTGAGCAAGCAACGCGAACTCAGCAGACCGGTTGGCG 910
QY 2023 GAGCAGACAGTAATACAGTCCCAATGTTGCGACCTAAGGTAAATGGGTAGATTACG 2082
DB 909 GAGCAGACAGTAATACAGTCCCAATGTTGCGACCTAAGGTAAATGGGTAGATTACG 850
QY 2083 CACTTTGAGCAGCGCCAGACGCCAGAGAGGAAACCATTAAGCTGTTGTTGCTCAAGATGA 2142
DB 849 CACTTTGAGCAGCGCCAGACGCCAGAGAGGAAACCATTAAGCTGTTGTTGCTCAAGATGA 790
QY 2143 TTGAATTAATACCGTTAACCTGTAAGCAATACCGTTAGTGGGCTTGGCCAAAT 2202
DB 789 TTGAATTAATACCGTTAACCTGTAAGCAATACCGTTAGTGGGCTTGGCCAAAT 730
QY 2203 CGCATTTTGTAGATTTTACGTGGCGGCGCAACCAATCACTTCAATCACTTTGGCCAA 2262
DB 729 CGCATTTTGTAGATTTTACGTGGCGGCGCAACCAATCACTTCAATCACTTTGGCCAA 670
QY 2263 CACCGTAATGCGACGAAATGCTGCAATACCAATACGATAGTGGCAACACCGAA 2322
DB 669 CACCGTAATGCGACGAAATGCTGCAATACCAATACGATAGTGGCAACACCGAA 610
QY 2323 AGTATTACCAATACCGAAAGCTGGGTTTCAAGCCTGTGATGACAGGTTTGGTAATCG 2382
DB 609 AGTATTACCAATACCGAAAGCTGGGTTTCAAGCCTGTGATGACAGGTTTGGTAATCG 550
QY 2383 TTATATAAATCAAAACCTTTCATGTCATATGCGCAGTGGCAAGACGAGGCAACGAG 2442
DB 549 TTATATAAATCAAAACCTTTCATGTCATATGCGCAGTGGCAAGACGAGGCAACGAG 490
QY 2443 CCAGTGTGAGACTTTCTCTTAATGCAAAATCATATTTCTAGAGGTAGGAATCGACAC 2502
DB 489 CCAGTGTGAGACTTTCTCTTAATGCAAAATCATATTTCTAGAGGTAGGAATCGACAC 430
QY 2503 CGGAGAGAGATTGAAAAATCGTGGATGGTGGCACTGGGGAAGACGCGGAAACCTTTC 2562
DB 429 CGGAGAGAGATTGAAAAATCGTGGATGGTGGCACTGGGGAAGACGCGGAAACCTTTC 370
QY 2563 TTATACGCGCATATGCGCAAGGCTGACGAGAGATTTGCTACAGTATGATCAATATTGGC 2622
DB 369 TTATACGCGCATATGCGCAAGGCTGACGAGAGATTTGCTACAGTATGATCAATATTGGC 310

QY 2623 ACCCATCACATAGAGATGCGGTTTCAACCGGTAAACCCAGGCAACGAGACCAACAT 2682
DB 309 ACCCATCACATAGAGATGCGGTTTCAACCGGTAAACCCAGGCAACGAGACCAACAT 250
QY 2683 AATAGAGTCACCGCTGCTTGAAGATTTAGTGCCTTGGCTCAATACCAATCATCA 2742
DB 249 AATAGAGTCACCGCTGCTTGAAGATTTAGTGCCTTGGCTCAATACCAATCATCA 190
QY 2743 TCTGTCAATGGGTGGGAAGCAATTCAAATAGAACTTTGCTGATGCGCGTTGCCA 2802
DB 189 TCTGTCAATGGGTGGGAAGCAATTCAAATAGAACTTTGCTGATGCGCGTTGCCA 130
QY 2803 TTTAAACCGCTGCCGACCATGCGACTGCAAGATGTAATATACGATGAAGCCAA 2862
DB 129 TTTAAACCGCTGCCGACCATGCGACTGCAAGATGTAATATACGATGAAGCCAA 70
QY 2863 GTTTCGCAAGCTGAGCCTTTCGTTGATGAGAAATGCGGCTGAG 2909
DB 69 GTTTCGCAAGCTGAGCCTTTCGTTGATGAGAAATGCGGCTGAG 23

RESULT 2

US-10-282-122A-40936
Sequence 40936, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykard, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPI7A.034A
CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 40936

LENGTH: 852

TYPE: DNA

ORGANISM: *Vibrio cholerae*

Query Match 29.0%; Score 842.4; DB 13; Length 852;
Best Local Similarity 99.3%; Pred. No. 1,26-267;

Matches 846; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 839 GTGAAACAGATTTAGATCTTTGTACAGCGCATGTCGATCAAGTGTGGTTGAAAT 898
Db 1 GTGAGACAGATTTAGATCTTTGTACAGCGCATGTCGATCAAGTGTGGTTGAAAT 60

QY 899 GAAGCAACGGGCAAGCGTTGTTGACGTGTATTAAGCCCATTTGACCTACATGTGGGC 958
Db 61 GAACGAACGGGCAAGCGTTGTTGACGTGTATTAAGCCCATTTGACCTACATGTGGGC 120

QY 959 AACCAATGATTTCTCTAGTACACTACAGCAAGATTTTGAAGCGCGTAGCCGAG 1018
Db 121 AACCAATGATTTCTCTAGTACACTACAGCAAGATTTTGAAGCGCGCGTAGCCGAG 180

QY 1019 TTGCTCGGCTATATTCGTGTTACGATATATGCGCGGATTTTGGCCATTTAGTACCAA 1078
Db 181 TTGCTCGGCTATATTCGTGTTACGATATATGCGCGGATTTTGGCCATTTAGTACCAA 240

QY 1079 ACCGCGGATGCTATGCGCAATTTAAACAGCAATGGCTCAACATCCTTACCGTAAAGGT 1138
Db 241 ACCGCGGATGCTATGCGCAATTTAAACAGCAATGGCTCAACATCCTTACCGTAAAGGT 300

QY 1139 GAGGATGACATGGGACCGGTGTATGTGTTCAGGGTATAGCTTGGGCTTAAGCCTGATGGT 1198
Db 301 GAGGATGACATGGGACCGGTGTATGTGTTCAGGGTATAGCTTGGGCTTAAGCCTGATGGT 360

QY 1199 GGTGATTTGACCAAGTTGAAAGATTTGATGATTTGAGCCGTGGCTTGTATGACCA 1258
Db 361 GGTGATTTGACCAAGTTGAAAGATTTGATGATTTGAGCCGTGGCTTGTATGACCA 420

QY 1259 GGTGAAATTTCTAATCTTCAATCCCGGATGAAATTTTCAATGGGATGTTTGGCCCTTGC 1318
Db 421 GGTGAAATTTCTAATCTTCAATCCCGGATGAAATTTTCAATGGGATGTTTGGCCCTTGC 480

QY 1319 ATGTACAGCCATCTTTTCAATGCTGCGGATACCTTGTATCTCAACAGTACTGACCGT 1378
Db 481 ATGTACAGCCATCTTTTCAATGCTGCGGATACCTTGTATCTCAACAGTACTGACCGT 540

QY 1379 TCATGATGATGCGCCTTGGGATTTGAATTTCAACATGGTGTGAGGTTATGTCCTTGGC 1438
Db 541 TCATGATGATGCGCCTTGGGATTTGAATTTCAACATGGTGTGAGGTTATGTCCTTGGC 600

QY 1439 CTGATGCGACAGATCAAGGAAAGCCGCGCTTGGCGTATCAACAGATGCTCAATGCG 1498
Db 601 CTGATGCGACAGATCAAGGAAAGCCGCGCTTGGCGTATCAACAGATGCTCAATGCG 660

QY 1499 CACATTTACCAAGATCAATCGAATTTGATGCGGATGTCGAGCTTAAACGTCGACCATTC 1558
Db 661 CACATTTACCAAGATCAATCGAATTTGATGCGGATGTCGAGCTTAAACGTCGACCATTC 720

QY 1559 CCAAGCGCTCAGTTCATCAATCCAAAGATTTAAACACTGCAAGATTTGAAACTGTG 1618
Db 721 CCAAGCGCTCAGTTCATCAATCCAAAGATTTAAACACTGCAAGATTTGAAACTGTG 780

QY 1619 GTCACTTTGATGATTTTGAAGTCAAGGATATCAAGTTCACAGATCTTATGATACCGG 1678
Db 781 GTCACTTTGATGATTTTGAAGTCAAGGATATCAAGTTCACAGATCTTATGATACCGG 840

QY 1679 TTTTCACTTAA 1690
Db 841 TTTTCACTTAA 852

RESULT 3
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US20030099277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: P3186P1
; CURRENT APPLICATION NUMBER: US/10/329,960

;; CURRENT FILING DATE: 2003-01-02
;; PRIOR APPLICATION NUMBER: US 09/643,990
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: US 08/487,429
;; PRIOR FILING DATE: 1995-06-07
;; PRIOR APPLICATION NUMBER: US 08/426,787
;; PRIOR FILING DATE: 1995-04-21
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 1830121
;; TYPE: DNA
;; ORGANISM: Haemophilus influenzae
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (4747)..(4747)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (9921)..(9921)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (10150)..(10150)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (29258)..(29298)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (36543)..(36543)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (36551)..(36551)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (40808)..(40810)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (44416)..(44416)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (44905)..(44905)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (44975)..(44975)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (45593)..(45593)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (45732)..(45732)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (47036)..(47036)
;; OTHER INFORMATION: n equals a, t, g or c
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (51334)..(51334)
;; OTHER INFORMATION: n equals a, t, g or c

```

NAME/KEY: misc_feature
LOCATION: (131340) .. (131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) .. (131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) .. (139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) .. (140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) .. (142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) .. (145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) .. (145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) .. (145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) .. (147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) .. (150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) .. (152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530) .. (152530)

```

	Query Match	25.6%	Score 744.2	DB 15	Length 1830121
	Best Local Similarity	66.7%	Pred. No. 1.4e-232		
	Matches 1134	Conservative	0	Mismatches 358	Indels 27
					Gaps 4
QY	2	AGAGGTTTGTATGCTCAGGGTTATCGACGTTTCCCAATATGACCCCGATTGTTT	61		
Db	958086	ATPAGAAATATATGATTAATCAATATTTATCTTCTCCCACTTTGATCCGATATTTT	958145		
QY	62	TCGATGGC-----CTCTACGGGTGGCGGATAGGCTTGATATTTGGGGGTTTC	115		
Db	958146	ACGCTTGGGATGATPATATGCGCTTAGCTTGATGCTTGATGATACCTTTAGGTTT	958205		
QY	116	CTTTTGTATGTGTTGGCCATGCGCCAGCGGATGCGGGCAGTGTGGACCGT	175		
Db	958206	GTTTTGCACGTGGCTTGCGGTTGCGGCTCTAATGCGCCAAATACGAGTTGGACAGTA	958265		
QY	176	GACGAGTCTCGACTGTATGTGCGCGGCTTTTAGTGTAGTGAATCGGTGCGCAGT	235		
Db	958266	GATCAAGTTGATAGCTTACTTTTCAACGGTTTATAGGGGGGTTATTTGGCGACGTTT	958325		
QY	236	GATTATGATCTCTACAAATTTTATGTCTTGGTGAACCCCTTTATTTATTCAA	295		
Db	958326	GGGAGTGTATTTTTCATATATCGATCATTTCTTACAAAGAACCACTTTATTTATTCGC	958385		
QY	296	GTTGACATGCGCGCATGTCTTCCACGCGGCTTATTTGGTGTGATCACCGGCATGTTC	355		

Db 958386 GTTGGAGAGTGGATGTGCTTCACGCTGAGCTTAATGATGTAATTTGGCATGAT 958445
QY 356 TGGATATGGCGCTAAACCAACGACCTTCTTGGTGTGGCCATTTTGTGGCCCTT 415
Db 958446 TGGACATCTTAATCTCAAAACGTAATTTTGGCAACGGCTAATTTTGTGGCCCTT 958505
QY 416 GTGCATTCGGTTTGGGATGGGACGTAATGTAATTTTGAATAGTGAATTTGGGGA 475
Db 958506 ATTCCGTTGGTTAGTTTGAAGAGATGTGTAATTTCTTAATCTTGAACATTTGGGA 958565
QY 476 CGAGTAACGATGTGCTTGGGCTTTGTAATCCCTTAATGTGGCC--CACTGGCCGC 532
Db 958566 CGCGAAACGAAATGCTTGGGCAATATTTCCCGAATATCTCTTTTACGCTCGT 958625
QY 533 CATCCCTACAGCTTTTGAATTCGCTTGAAGGCGGTTGCTTCTTATCTTAAT 592
Db 958626 CATCATACACACTTTTGAAGCTTTTGAAGGCTTGTGTTGTTTACGATTTGAA 958685
QY 593 TGGTTATGCTAAACCTTCGCTGAGGACGCTATCCGACTTTTATGCTGATAC 652
Db 958686 ATTTTATTAATAAACACGCTCAATGCTTGTGACAGTTTATCTTAATGCTTAT 958745
QY 653 GGTACATTCGCTTCCTTGGGATACGTCGAGGACAGATGCTCACTGGGCTGTT 712
Db 958746 GGGCTCTTCGTTTATTTTGGAAATATGCTGTAACCTTAAG-----TTGAAATTT 958799
QY 713 GGTGGCTTCAATTCATAGGCGCAATCTCTCTTACCTTATGATCATCGGTAATTT 772
Db 958800 TTTGGATTAATTAACAGAGGCAAGCCCTTGTCTGCGCATGATATGTTGGTGTCTTC 958859
QY 773 ATGATGTTGGTCTTCAAGCGGCTTGTATCAACCGTGTAGCAGAAATTTGGGT 832
Db 958860 ATTATGCTTGGCTTATTCACGC-----AAAGTGGCGCATTAATATGAG 958907
QY 833 AGTAGGGAACAGTATTTAGATCTTGTACGCGATCGCATCAAGGTTGTTGGGT 892
Db 958908 ATTTTATGAAGCAATCTTGAAGCTTGTGCGCATGTTTACTGAAAGGGAATGGGT 958967
QY 893 GAAATTAACGACGCGCAAGCGTGTGTAATGATGTAATGCGGATTTGACCTAGAT 952
Db 958968 GCTATATACGCTACAGGTAAGCATTTGCTCAAGGTCATTAATGATATGAT 959027
QY 953 GTGGGCAACATCGTTTCCCTAGTACTACACGAGAGTTTGGAAAGCTGCGCTA 1012
Db 959028 GTGGCAATATCAATTTCCGCTGATTAACCGTAAGTAATTTGAAAGGCGGAT 959087
QY 1013 GCCGAGTTGCTGGCTATTTCTGTGTAGCATATGCGCGGATTTTCCCAATTAGT 1072
Db 959088 GCTGAATTTTAAAGTTATATCGTGAATGACATGCCGTGATTTCCGCGACTGGC 959147
QY 1073 ACCAAACCTGGGATGCTAATGCCAATTTAAACCAAGATGGCTCAACAATCTTACCGT 1132
Db 959148 ACGAAACCTGGGATGCTAATGCCAATTTAAACCAAGATGGCTTGCATTCGCT 959207
QY 1133 AAAGTAGAGTACATGAGGACGCGTATGATGTTTCAAGGTAAGCTTGGCTAAAGCT 1192
Db 959208 AGAGCGTGAATGATATGCGCGCGTATATGCTGCAAGGCAAGATGCGCTAAGCT 959267
QY 1193 GATGCTGCTATATTTGAACAGTTGAAAAAGATTTGTATGATTTAGCCGTGCGCTT 1252
Db 959268 AATGAGAAACTATGATGATGCTACGTAATTTGTAATTAATTAAGAAAGGATTTGAT 959327
QY 1253 GACGAGGTGAATTTCTTAATCTTACATCCGCGTAATTTTCAATGAGGCTGTTGGCG 1312
Db 959328 GATAGAGAGAGATTTTAATCTTTTAACTTGGGAAATTTGATCTTGGTGTCTTCGT 959387
QY 1313 CTTTGCATGTAACGACCATTTTTCATGCTGAGGAGATACTTGTATCTCAACAGTACT 1372
Db 959388 CCTTGTATGATGACGATACCTTTTCTCTTGTGGCGATTTTCACTTACATGATAT 959447
QY 1373 CAGGCTTCAATGATGAGCCTTGGGTTGAAATTTCAACATGGGAGGTTTATGTTG 1432
Db 959448 CAGGCTTCAATGATGAGCCTTGGGTTGATGATTAATTAATCAAAATTCAGGTTTACCTTC 959507

QY 1433 CTGGCGTGAATGGCAGATCAAGGAAAAAGCCGGCTTGGCGTATCAAGATGCTC 1492
Db 959508 TTAGCATTAATGCTTCAGATCAAGGCAAAAAAGCCGCAAGCATATCAATTAATTTGTG 959567
QY 1493 AATGCGACATTTTACCAAGATCACTCGAATTTGATGCGGATGTGCACTTAAACGTGAG 1552
Db 959568 AATGCGCATATTTTAAAGATACGCTTGAATATGCGTGAAGTACATTTAAACGCGAG 959627
QY 1553 CCATTCACAGCCCTCAGTTCATATCAATCCAAAGATTTAAACCTGACAGATTTGGAA 1612
Db 959628 CCTTCCCATTAACCAAAATTAAGAAATTAATCCAGATTAATAAAACCTTGAAGATTTGGA 959687
QY 1613 ACTTGGGTCATTTGATGATTTTGGACGTCACCGGATATCAATTGCTCAGATCTTATGAA 1672
Db 959688 ACTTGGGTCAGATGATGATTTTAAAGTCGTTGCTATCAATCCACAGAACCAATTAA 959747
QY 1673 TACCGTTTCACTTAAT 1691
Db 959748 TATCCTTTTCCGCTTAAT 959766

RESULT 4
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US2004018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Hemophilus influenzae Rd Genome, Frak
; FILE REFERENCE: P8186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; PRIOR FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

```



```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 30692
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-282-122A-30692

Query Match      17.4%; Score 505.4; DB 13; Length 852;
Best Local Similarity 74.6%; Fred. No. 7.9e-156;
Matches 655; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY      840 TGAACAGATTTAGATCTTGTGACGCGATCGTGCATGCAAGTGTGTGGTTGAATG 899
DB      2 TGAACACATTTAGATCTTGTGACGCGATCGTGCATGCAAGTGTGTGGTTGAATG 61

QY      900 AACGAAACGCGCAAGCTGTGTGACGCGATCGTGCATGCAAGTGTGTGGTTGAATG 959
DB      62 CAGGACGCGCAAGCTGTGTGACGCGATCGTGCATGCAAGTGTGTGGTTGAATG 121

QY      960 AACATCAGTTCTCTAGTGAATGCAAGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 1019
DB      122 ATAAACCAATTCGCTGATGACCAAGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 181

QY      1020 TCGTGGCTATATTCGTGTGATGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 1079
DB      182 TTTTAGGTATATTCGTGTGATGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 241

QY      1080 CCGGAGATGCTATGCAATTTAAACCAAGTGTGTGGTTGAATGCGGATGCGGAGT 1139
DB      242 CCGGAGATGCTATGCAATTTAAACCAAGTGTGTGGTTGAATGCGGATGCGGAGT 301

QY      1140 AAGATGACATGCGGAGTGTGTGATGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 1199
DB      302 CCGATGATGCGGAGTGTGTGATGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 361

QY      1200 GTCATATGACCAATTTAAACCAAGTGTGTGGTTGAATGCGGATGCGGAGT 1259
DB      362 AAACGTCGATCAATTTGCGCAAAATCGCATATTCGCAACGCGATGATGACCGTG 421

QY      1260 GGAATTTCTTAATCTTCAACATCGGATGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 1319
DB      422 GGAATTTCTTAATCTTCAACATCGGATGCAAGTGTGTGGTTGAATGCGGATGCGGAGT 481

QY      1330 TGTACAGCATATTTTCAATGCTGTGGGATTAATCTTCAACAGTACTACAGCTT 1379
DB      482 TGTACAGCATATTTTCAATGCTGTGGGATTAATCTTCAACAGTACTACAGCTT 541

QY      1380 CATGATGCTGCTGTGGGATTAATCTTCAACAGTACTACAGCTT 1439
DB      CATGATGCTGCTGTGGGATTAATCTTCAACAGTACTACAGCTT 1439
```

```

DB      542 CTGTGTATGTCGCGCTTGCTGATTTCAACCAATTTAAGTGTTCACCTTTCTTGCT 601
QY      1440 TGATGCAACATCAACAGGAAAAAGCCGCGCTTGCTGATTCAGAGATGCTGCAATGCGC 1499
DB      602 TGATGCAACATTTAGCGGCAAAAAAGCGGTAAGCGGTACCACAAAATCAATGACAC 661
QY      1500 ACATTTCAAGATCAACTGCAATTTGATGCGGATGCAAGTGTGAATGAAGCGGATGCGC 1559
DB      662 ACATTTCAAGATCAACTGCAATTTGATGCGGATGCAAGTGTGAATGAAGCGGATGCGC 721
QY      1560 CAGGCGCTCAATTCATCAATCAATCAAGATTTAAACACTGCAAGATTTGAAACTTGGG 1619
DB      722 CGTTCACCAATTTACATTTAAACCGGATTTAAACCTCTCCAGATTTAGAACTTGGG 781
QY      1620 TCACTTGAATGATTTTGAAGTGTGACGCGATTCAGTTCAGATTCCTTAATACCTGCT 1679
DB      782 TCACTTGAATGATTTTGAAGTGTGACGCGATTCAGTTCAGATTCCTTAATACCTGCT 841
QY      1680 TTTGAGTCTAA 1690
DB      842 TCTCTGTGTAA 852

RESULT 6
US-10-282-122A-22162
; Sequence 22162, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPI/TA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 22162
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-282-122A-22162
```

Query Match	15.9%;	Score 462.2;	DB 13;	Length 852;
Best Local Similarity	71.4%;	Pred. No. 1.7e-141;		
Matches 608; Conservative	0;	Mismatches 243;	Indels 0;	Gaps 0;

QY	840	TGAAACAGATTTAGATCTTTGGCAGCGCATCGATCAAGGTGTTGGGTGAAGT	899
Db	2	TGAGCAATATCTGAGCTTTGTCCGCCGATGTTAGTGAAGGGAATGGTTCCTAATG	61
QY	900	AACGAAACGGGCAAGCGTTGTTGACTGTGATTAATGCCGATTTGAACCTACGATGTGGCA	959
Db	62	AACGTACAGGTAGAGCAATTCCTCCACCGGTCAATTAATGCAATTTAGAAATATGATGTGGGGA	121
QY	960	ACAATCAGTTTCCCTCAGTGAACCTACACGCAAGGTTTTGGAAAGTGCAGAGCGAGT	1015
Db	122	AATATCAATTTCCGCTGATTTACTACCCGTTAAAGTATTTGGAAAGCGCGATTTGCTGAAT	181
QY	1020	TGCTGGGCTAATTTGTGTGTTACGATATGCGGCGGATTTTGGCAATTAAGTACCAAA	1079
Db	182	TTTTAGTTAATGTGTGATATGCAAGCCGCTGATTTCCGCCACCTTGGCAGAA	241
QY	1080	CTTGGAATGCTAATGCAATTTAAACCAAGCATGGCTCAACAATCCTTAACGTTAAAGGTG	1139
Db	242	CTTGGGATGCTAATGCGAATGAAATGACGCTTGCTTGCAATCCGATGTGTAAGCG	301
QY	1140	AGATATCACATGGGACGCGTGTATGTGTTCAGGGTACAGCTTGAGCCTGATGTGT	1199
Db	302	TTGATGATATGTGGGCGCGTATATGTGTGTCAGGACAGACATGTGCGTAAGCCTAATGAG	361
QY	1200	GTCATATTGACCAAGTTGAAAAAGATTTGTGTGATGTTAGACCGTGGCCTTATATACGAG	1259
Db	362	AAACATGATGATAGCTACGTAAATTTTATTAATCTTAACGAAAGTATATGATGTAG	421
QY	1260	GTGAATTTCTTAATCTTACAAATCCGGGTGAATTTACATGGGGTGTGTGCGCCTTGCA	1319
Db	422	GAGAGATTTTAATCTTTTTTAACCTCGGGAAATTTGATCTTGTTGTCTTGCTCTTGA	481
QY	1320	TGTACAGGCATATTTTTCATTTGCTGGGGGAACTCTGTATCTCAACAGTACTAGCGTT	1379
Db	482	TGCATACGATCTTTTCTTCTTCTTGGGCGATCTTAACATCTTAATGATACGCGTT	541
QY	1380	CATGTGATGTGCCCTTGGGGTGAATTTCAACATGTCAGGTTATGTGTCTTGTCGC	1439
Db	542	CCTGTGATGTGCCCTTGGAATTTGAATTTCAATCAAAATTCAGGTGTTTACCTTTTAGAC	601
QY	1440	TGATGGCACAGATCAACAGGAAAAACCGGCGTTGGCCTATCACAAATCGTCAATGCG	1499
Db	602	TTATGCTCTAGATCAAGGCAAAAAAGCCGGCAAGCATATCATAAATTTGATGATGCG	661
QY	1500	ACATTTACCAAGATCACTCGAATTTGATGCGGATGTGCAGACTTAAACGTGAGCATTTCC	1559
Db	662	ATATTTATGAAGATCAGCTTAGATTTAATGCCGACGTACCACTTAAAGCGAGCCTTTC	721
QY	1560	CAGGCGCTCAGTCCATATCAATCCAAAGATTTAAACACGTGACGATTTGAAACTTGGG	1619
Db	722	CATTAACAAATTTAGAAATTTATTCAGATATTAATAAACCTTGAAATTTGAAACTTGGG	781
QY	1620	TCACTTTGAATTTTGAACGTCAACCGGATATCAGTTTCAAGATCCTAATCAATACCGGT	1679
Db	782	TCACGATGATGATTTTAAAGCGTGTGCTATCAATCCACGAAACCAATTAATTAATCCTT	841
QY	1680	TTTCAGTCTTA 1690	
Db	842	TTTCGCTTAA 852	
RESULT 7			
US-10-282-122A-32325			
: Sequence 32325, Application US/10282122A			
: Publication No. US20040029129A1			
: GENERAL INFORMATION:			
: APPLICANT: Wang, Liangsu			
: APPLICANT: Zamudio, Carlos			
: APPLICANT: Malone, Cheryl			

```

1  / APPLICANT: Haselbeck, Robert
2  / APPLICANT: Ohlsen, Kari
3  / APPLICANT: Zyskind, Judith
4  / APPLICANT: Wall, Daniel
5  / APPLICANT: Trawick, John
6  / APPLICANT: Carr, Grant
7  / APPLICANT: Yamamoto, Robert
8  / APPLICANT: Forsyth, R.
9  / APPLICANT: Xu, H.
10 / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
11 / FILE REFERENCE: ELITRA 034A
12 / CURRENT APPLICATION NUMBER: US/10/282,122A
13 / CURRENT FILING DATE: 2003-02-20
14 / PRIOR APPLICATION NUMBER: 60/191,078
15 / PRIOR FILING DATE: 2000-03-21
16 / PRIOR APPLICATION NUMBER: 60/206,848
17 / PRIOR FILING DATE: 2000-05-23
18 / PRIOR APPLICATION NUMBER: 60/207,727
19 / PRIOR FILING DATE: 2000-05-26
20 / PRIOR APPLICATION NUMBER: 60/230,335
21 / PRIOR FILING DATE: 2000-09-06
22 / PRIOR APPLICATION NUMBER: 60/230,347
23 / PRIOR FILING DATE: 2000-09-09
24 / PRIOR APPLICATION NUMBER: 60/242,578
25 / PRIOR FILING DATE: 2000-10-23
26 / PRIOR APPLICATION NUMBER: 60/253,625
27 / PRIOR FILING DATE: 2000-11-27
28 / PRIOR APPLICATION NUMBER: 60/257,931
29 / PRIOR FILING DATE: 2000-12-22
30 / PRIOR APPLICATION NUMBER: 60/267,636
31 / PRIOR FILING DATE: 2001-02-09
32 / PRIOR APPLICATION NUMBER: 60/269,308
33 / PRIOR FILING DATE: 2001-02-16
34 / Remaining Prior Application data removed - See File Wrapper or PALM.
35 / NUMBER OF SEQ ID NOS: 78614
36 / SOFTWARE: PatentIn version 3.1
37 / SEQ ID NO 33325

```

Query Match	13.9%	Score 403.8;	DB 13;	Length 849;
Best Local Similarity	67.3%	Pred. No. 4e-122;		
Matches 570;	Conservative 0;	Mismatches 277;	Indels 0;	Gaps 0;

QY	840	TGMAACAGATTTAAATCTTTGTCAAGCGCATCGTCATCAAGGTGTTGGGTTGAAATG	899
Db	2	TGAGCAGATATCTGGCATGTGTCAAGCATTATCGATGAAGGTGATGATTGATTAATA	61
QY	900	AACGAAACGGGCAAGCGTTGTTGACTGATTAATGCCAATTTGACCTACGATGATGAGGA	959
Db	62	AACGACAGGAACCCGTTGTTTAAACGTGATTAATGCCATCTGGAATATGATGTTGCCA	121
QY	960	ACATTCAGTTCTCTAGTGACTACACGCAAGATTTTGGAAAGCTGCCGATGCCAGT	101
Db	122	ATACCAATTTCCATCTCATACAAACGGTAAAGTTTTTCAAAAGCGGATTTGGCAGAC	181
QY	1020	TGCTGGCTATATTTGTGTGTTACGATATATGCGCGCGATTTTGCCCAATTAGTACCAAA	107
Db	182	TATTAGGGTACTTAAGTGTATATATATACAGACAAATTTCTGTCTATTGGCTGTAAATA	241
QY	1080	CCGTGGATGCTATATGCCATTTTAACAACAGATATGGCTCAACATCTTACCGTAAGAATG	113
Db	242	CGTGAACGCCAATCTATATGAAATAATGCCATGTGTTAATTAATCCGCATCGTAAGGGTG	301
QY	1140	AGCATGACATGGGACGCGGTATGATGATTAAGGATGAGAGCTTGGGCTAAAGCCTGATGTG	1199
Db	302	AAGATGATATGGGCGAGATCTATGTGTCAAGGACCCCAATGCAACGCCCTGATGTCT	361
QY	1200	GTCAATATGACCACTGTGAAAAAGTTGTGATATTTGAGCCGTGGCTTGATGACCGAG	1259
Db	362	CGCACTTTGATCAATATCGTAAGGTGTAAATTAATCTAAACGGTATTTGACATCTGG	421

QY 1260 GTGAATTTCTTAATCTTCAAAATCCGGGTGAATTTACATGCGGGTGTGGCCCTTGA 1319
DB 422 GTGAGATGACTCTTTTATATATCCCGAGAAACAGCGTTAGGGTGTTCAGTCCATGCA 481
QY 1320 TGTACAGCATATTTTCAATGCTGCGGGATACCTTGTATCTCAACAGTACAGCGTT 1379
DB 482 TGTATACATATCTTTTTCATGCTGCGGTGACAGGCTTATTTAACCTCAATCAAGTA 541
QY 1380 CATGTGATGTCCCTTGGGGTGAATTTCAACATGCTGAGGTTTGTGTTCTTGGCG 1439
DB 542 GCTGTGATGTCCCATTAAGGTGAATTTTCAATTCATGCTTGTATTTATTTACCGC 601
QY 1440 TGTATGCAAGATCAAGGAGAAAGCGCGCTTGGCGTATCAAGATGCTCAATGCGC 1499
DB 602 TGTGTGCTCAATTCAGAGGCAATTAACAGGTAAGATTCATTAAGATGTAATGCTC 661
QY 1500 ACATTTTCCAGATCACTCGAATGATGCGCGATGCGAGCTTAAACGTGAGCCATTC 1559
DB 662 ACATTTATGAGAACCAATTAACCTGATGCGTGAATGTTCACTTAAAGAGAGCCATTC 721
QY 1560 CAGCGCTTCAGTTCATATCAATCCAAAGATTTAAACAGTCAAGATTTGGAAACTTGG 1619
DB 722 CATTCCTAGCTATATATTAATCCAAATTAAGATGATGATTAAGAACTTGG 781
QY 1620 TCACTTTGATATTTTGAAGCTACCGGATATCACTTCCAGATCTTATTAATACCGT 1679
DB 782 TCAGCAGATGACTTGTAGTGTAAAGGTATCAGTGTACAGAGCGATTAATACCGT 841
QY 1680 TTTTCACT 1686
DB 842 TCAGGCT 848

RESULT 8

US-09-741-669-182
Sequence 182, Application US/09741669
Patent No. US2002022718A1
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 182
LENGTH: 876
TYPE: DNA
ORGANISM: *Bacterichia coli*
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(876)
US-09-741-669-182

Query Match 8.5%; Score 248.2; DB 9; Length 876;
Best Local Similarity 58.8%; Pred. No. 1.7e-70;
Matches 510; Conservative 0; Mismatches 288; Indels 69; Gaps 2;

QY 14 ATGCTCAGGTTATCTGCAATTTCCCAATTTACCCCGTATGTTTGTGATGCGCCCT 73
DB 1 ATGACCACTACCTATCTGCAATTTCCCGAGTTGATCCGATTTCTCAATTAAGACCC 60
QY 74 CTAGCGTGCCTGCTGATGCTGATGATTTGTTGGGTTCTTTTGTATGCTG 133
DB 61 GTGGGCTTCACTGTTAGGCGCTGATGATGCTGCTTCAATTTTGAATGTGCTG 120
QY 134 GCATATCCGCGAGGCGATCCGCGGCGAGTGTGAGCGGTAGCAGAGTCTGTGCTG 193

DB 121 GCAACGACGCGGGAATCGTCCGGCGAGCGCTGACCCAAATATAGATTGAAACTTA 180
QY 194 TTTATGCGCGCTTTTATGAGTATGATGATGAGGCGAGATTGATGATCTTAC 253
DB 181 CTATATGCGGCTTCTCGGCGCTCTTCTCGGCGGAGAGATGATGATGCTGCTGAC 240
QY 254 AATTTGATGCTGCTGAGACCCCTTATTTATTAAGAGTGGACGCGGCGATG 313
DB 241 AATTTCCGAGTTATGCGGAGTCCGATGATGCTGCTGCTGAGAGCGGCGATG 300
QY 314 TCTTTCACGCGGCTTATGAGTGTGATGATCCGCGCATGTTGTGATGCGGTAATAAC 373
DB 301 TCTTTCACGCGGCTGATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 360
QY 374 CAAGCAGCTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 433
DB 361 AAGCTTCTTCTTCCAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 434 ATGGAAGTATGCGTATCTTATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 421 GCGGCGGCTGCGGCACTTATTAAGGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 -----AAGGAT 487
DB 481 CCGTTGCGCATGCTTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 540
QY 488 GTGCTTGGGCTTTTGTATTCCTTAATGAGGCGCATCGCGGCGCATCTTACAGCTT 547
DB 541 CCGCAGTGGCAATTCATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 600
QY 548 TATGATTCGCTTAGAGGCGTGTGCTGCTTCTTATTTATTTATTTATTTATTTATTT 607
DB 601 TACAGCTGCTGCTGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 608 CCGTGCCTGAGCAGCGTATCCGAGCTGTTTATGCTGATGATGATGATGATGATGATG 667
DB 661 CCAGCCCAATGAGGAGTGTCTCAGGTTTGTCTATGTTATGCTTACGCGCGTTTGGCATC 720
QY 668 CTGTGGAATACGCTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCGTGTGAGGCG 724
DB 721 ATTTGATGTTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 725 TCAATGCGGCAATCTCTCTTACCTATGATGATGATGATGATGATGATGATGATGATG 784
DB 781 AGCATGCGGCAATCTCTTCCATCCGATGATGATGATGATGATGATGATGATGATGATG 840
QY 785 TCTTACAGCGGCTTGTATCAAGAC 811
DB 841 GCATATGCTGCGAGCCGACAGCAAC 867

RESULT 9

US-09-790-988-1
Sequence 1, Application US/09790988
Patent No. US20020127687A1

GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEMI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentia Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: *Buchnera sp.*

US-09-790-988-1

Query Match 5.2%; Score 152.2; DB 9; Length 640681;
Best Local Similarity 52.6%; Pred. No. 1.7e-16;
Matches 433; Conservative 0; Mismatches 323; Indels 66; Gaps 2;

```
QY 35 TTTTCCCAATATGACCCCGTATGTTTGGATGCGCCCTTACCGGTGCGGTATGCG 94
DB 479304 TTTTCCCAATATGACCCCGTATGTTTGGATGCGCCCTTACCGGTGCGGTATGCG 94
QY 95 TTGATGATTTGGGGGTTTCCCTTTTGGATGCGGTGCGGTATGCGGTATGCGGTATGCG 154
DB 479364 TTTTATGATTTTAAAGTTTATTTTATTTGCTATGATGGAAGAAAGTACAGTATTA 479423
QY 155 GCGGGGAGTGTGGAAGCGGTGAGCAAGTCTGACTTGTATTCGCGCGCTTTTAAAGT 214
DB 479424 AATTAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479483
QY 215 GTATGATGCGGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 274
DB 479484 TCAATGATGCGGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479543
QY 275 GACCTCTTATTTATTTCAAGTGTGACTGCGGTATGCGGTATGCGGTATGCGGTATGCG 334
DB 479544 AATATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 479603
QY 335 GGTGTGATGACCGGTATGCGGTATGCGGTATGCGGTATGCGGTATGCGGTATGCGGTAT 394
DB 479604 GGTGTGATGACCGGTATGCGGTATGCGGTATGCGGTATGCGGTATGCGGTATGCGGTAT 479663
QY 395 GCGGATTTTGTGCGGTATGCGGTATGCGGTATGCGGTATGCGGTATGCGGTATGCGGTAT 454
DB 479664 TCGATTTTATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 479723
QY 455 ATGAATAGTGAACCTTTGGGACGAGT---AACGATGCGCTTGTGCTTGTATTCCT 511
DB 479724 ATTAATAGTGAACCTTTGGGACGAGT---AACGATGCGCTTGTGCTTGTATTCCT 479783
QY 512 AATGCT----- 517
DB 479784 AATTTCAAAATCAGATTTAAAGAAATTAATAATCCGAAATACACTATTAATA 479843
QY 518 -----GCGCCACTGCGCGCGCCATCCTTCAAGCTTATGAATTCGCTTAGAAGGC 568
DB 479844 GATCAATAGGAGCAATGCGCGCTATCTCACTCACTCACTCACTCACTCACTCACTCA 479903
QY 569 GTGCTTCTGTTCTTATTTCTTATTTGTTTATTTGTTTATTTGTTTATTTGTTTATTTG 628
DB 479904 ATTTCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 479963
QY 629 TCCGACTGTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
DB 479964 ACGGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 480023
QY 689 CCAATGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
DB 480024 CCAATGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
QY 749 CCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 480084 CCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
```

RESULT 10
US-10-672-787-35
; Sequence 35, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATERSON, Chandra
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELIPIRA.025C1

CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 09/596,002
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 96109
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-10-672-787-35

Query Match 4.9%; Score 142.2; DB 13; Length 96109;
Best Local Similarity 59.1%; Pred. No. 8.5e-34;
Matches 262; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

```
QY 38 CCAATATTTGACCCCGTATGTTTGCATGCGCCCTTACCGGTGCGGTATGCGGTATGCG 97
DB 51130 CCAAGATTTAACCTGTTGCTTTAGACTTGGGCAATGATGATGATGATGATGATGATGAT 51189
QY 98 ATGATTTGTTGCTTCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 157
DB 51190 ATGATTTGTTGCTTCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51246
QY 158 GCGAGTGTGAGCGGTGAGCAAGTCTGACTGATTTATTCGCGGTATTTAGATGATA 217
DB 51247 GCAACAGATTTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51306
QY 218 GTATGCGGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 277
DB 51307 ATTTAGTGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51366
QY 278 CCTTTATTTATTTCAAGTGTGACTGCGGTATGCTTCCAGCGGTATTTAGGT 337
DB 51367 CCTTTATTTATTTCAAGTGTGACTGCGGTATGCTTCCAGCGGTATTTAGGT 51426
QY 338 GTATGACCGGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 397
DB 51427 GTGTTGCTTGCAGTGTGATTTTGTGACATTAATAATAATAAAGCACTTTTACCGGTATG 51486
QY 398 GATTTGTGCGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 457
DB 51487 GATTTGTGCGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51546
QY 458 AATATGATGATTTGGGACGAGT 480
DB 51547 AATGCGAGCTTTGGGACGAGT 51569
```

RESULT 11
US-10-297-465A-1/c
; Sequence 1, Application US/10297465A
; Publication No. US20040142413A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Andrew
; APPLICANT: Reinach, Fernando
; APPLICANT: Seubal, Joao
; APPLICANT: Medians, Joao
; TITLE OF INVENTION: Isolated Genome of Xylella fastidiosa and Uses Thereof
; FILE REFERENCE: PABESP 202 US (10213376)
; CURRENT APPLICATION NUMBER: US/10/297,465A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01618
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,906
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2731748
; TYPE: DNA
; ORGANISM: Xylella fastidiosa

US-10-297-465A-1

Query Match 4.0%; Score 115.4; DB 17; Length 2731748;
Best Local Similarity 53.8%; Pred. No. 8e-24;
Matches 261; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 36 TTCCCAATATGACCCCGATATGTTTCGATCGCCCTCAGCGGTGCGGTATGCGCT 95
DB 2214488 TCACAGCTATCGATCGATGCTTCTTCTACTGTCGCGGTGAGAGTGCATGATGCGCT 2214429
QY 96 TGATGATATGTTGGTGGTTCCTTTTGTATGTTGGTGGCAATGCGCGATCGCG 155
DB 2214428 TGATGATATGTTGGCAAGTTTGGGCGCGGTGCTGTTGGAGCGCAATCGATCGAGCGG 2214369
QY 156 CGGAGAG---TGGTGGAGCGCGGTAGCAAGTCTGACCTGTTATTCGCGCGCTTTTGG 212
DB 2214368 GCGGTTGCTGTTGTTGTCACATGAGTGTGTTTTCGACCTGTTGTTACGCAATATG 2214309
QY 213 GTGATGATGCGTGGCGGAGTTGTTATGATCTTCTACATTTTGTATCTGTTGCTTG 272
DB 2214308 GTGATGATGCGTGGCGGAGTTGTTATGATCTTCTACATTTTGTATCTGTTGCTTG 2214249
QY 273 CTGACCTCTTATTTATTTCAAGTGTGAGTGGCGGATGCTGTTCCAGCGCGCTTAT 332
DB 2214248 AAGACCATGCTGCTGTTGTTGAGTGTGAGAGGGGCGCATGATCTTTCATGAGGTTGA 2214189
QY 333 TGGGTGTATGACCGCGCATGTTCTGTATGCGCGGTAAACCAACGACCTTCTTGGTG 392
DB 2214188 TGGGTGTATGACCGCGCATGTTCTGTATGCGCGGTAAACCAACGACCTTCTTGGTG 2214129
QY 393 TGGCGGATTTGTTGCCCCCTTATGTCATTCGTTGGGAGTGGAGCGATGCGTAT 452
DB 2214128 TGGCGGATTTGTTGCCCCCTTATGTCATTCGTTGGGAGTGGAGCGATGCGTAT 2214069
QY 453 TTATGATATGTAATTTTGGGAGAGTAAACGATGTCCTTGGGCTTTTGTATTCCTA 512
DB 2214068 TCATGCGCGGTGAGCTTGGGAGAGTAAACGATGTCCTTGGGCTTTTGTATTCCTA 2214009
QY 513 ATGCT 517
DB 2214008 GTGCT 2214004

RESULT 12

US-10-152-319A-1589/c
; Sequence 1589, Application US/10152319A
; Publication No. US20040072160A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT FILING DATE: US/10/152,319A
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928

;; PRIOR FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: US 60/330,867
;; PRIOR FILING DATE: 2001-11-01
;; PRIOR APPLICATION NUMBER: US 60/330,462
;; PRIOR FILING DATE: 2001-10-22
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 2221
;; SOFTWARE: Patent Ver. 2.1
;; SEQ ID NO 1589
;; LENGTH: 2440
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
;; OTHER INFORMATION: Genbank Accession No. NM_013030

US-10-152-319A-1589

Query Match 2.9%; Score 84; DB 12; Length 2440;
Best Local Similarity 58.5%; Pred. No. 1.1e-15;
Matches 166; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

QY 2480 TCTAGAGTGAAGATCAGACCGCGAGAAAGTTGAAAAATCGTGTGTTGCGACTG 2539
DB 766 TCAGGGGCGAGCGAGACCGAGAACAGACAGCGAGTTAAACAGTTCATGACTGTCCTT 707
QY 2540 CGGAAGCAGCGGAAACTTCTTTACAGCGCATATGSCCAAGCTGACGAGATATG 2599
DB 706 GCAAAACCGCGCTGAGTCAAGTCTGTCCTCCCGCTGATCAGGCGCAATGTTG 647
QY 2600 GTCACTAGTATCAATATTTGCAACCCATCACATAGAAATGCGG---TTTCAACCGT 2656
DB 646 GTAGCAGAGTCCCGATGTTGAGCCCATGATTTATGGAATGCGGAGTCACTTCAAC 587
QY 2657 AACCCACCGCGAGACGACCAATATGAAAGTCAACGCTGTTGAGATGATGCT 2716
DB 586 AAGCAGAGAGACCAATGCTGACAAATGATGAGTCTGAGCTGACCTGACCGAGA 527
QY 2717 GCGTTGCGCACTAAACCAATTCATCTCTGCAATTTGCGTGGGA 2760
DB 526 ACTGTACAGAGATTCACACCAACCAACCTGCGACTGGTTTGA 483

RESULT 13

US-10-152-319A-1988/c
; Sequence 1988, Application US/10152319A
; Publication No. US20040072160A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT FILING DATE: US/10/152,319A
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867

PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 1988
LENGTH: 3950
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURES:
OTHER INFORMATION: Genbank Accession No. NM_053380
US-10-152-319A-1988

Query Match 2.7%; Score 77.8; DB 12; Length 3950;
Best Local Similarity 56.7%; Pred. No. 1.9e-13;
Matches 164; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

QY 2480 TCTAGAGTGAAGATGACGACCGGAGAAAGATTGAAAAATCGTGATGGGCACTG 2539
DB 724 TCCAGAGTGAAGACCAACACAGAGGCGAGTGAAGAGTCATGACAGTAGCTCCT 665
QY 2540 GCGAAGCAGCGGGAACCTCTCTTTACAGGCGATATGCGCAAGCTGAGAGATATG 2539
DB 664 GCAATATCCCTTCTGAATCTATTTCTCTCTGCTGCTGCAATAGTCCCAATGCTG 605
QY 2600 GTCACTAGTACCAATATTGGCAACCCATCAGCAATGAAATCGCGTTTCAAC---CGT 2656
DB 604 GTATGAGAGTCCCGATGTTAGACCCCATGATGATGGGATGCGAGCGGACTGACAG 545
QY 2657 AACCCACCGGCAAGACCAACATATATGAAATGACCGCTGTTGAGATTGAATGAGT 2716
DB 544 AAGAGAGGCAACCATGCTGACGATATTGAGAGGACGCTGTGAGCTTTGACACAG 485
QY 2717 GCCCTGCCCAATAACCATATCATCTGCAATTTGGGGAGAGCA 2765
DB 484 ACTGTACACGACACCCCTATCAAGTCCAGCCACAGATTAACATTA 436

RESULT 14
US-10-276-774-1331/c
Sequence 1331, Application US/10276774
Publication No. US20040053245A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Teng, Y, Tom et al
TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-030
CURRENT APPLICATION NUMBER: US/10/276,774
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 2700
SOFTWARE: Custem
SEQ ID NO 1331
LENGTH: 2208
TYPE: DNA
ORGANISM: Homo sapiens
US-10-276-774-1331

Query Match 2.6%; Score 75.4; DB 13; Length 2208;
Best Local Similarity 56.0%; Pred. No. 7.6e-13;
Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGTGAAGATGACCGGAGAAAGATTGAAAAATCGTGATGGTGC 2535
DB 738 CACCTCCAGGCGGACAGACCAACAGGACGCGAGTGAAGAGTCAATGACAGTGC 679
QY 2536 ACTGGGAAAGCAACGCGGAACTCTTTACAGCGCATATGCGCAAGCTGACGAGT 2595
DB 678 TCCGTGAAAGCTCTTCTGTAATCTCCGATCTCCACCTGCGATGAGCGCAACATAGT 619

QY 2596 ATTGCTCAGATGATACCAATATTGGACCCATCCATAGGATATCGGCTTCAACCG- 2654
DB 618 GTTGGATGATGAGTTCATATGTTGGCCCCCATGATATGAGGAGCGCCGAAACGT 559
QY 2655 -GTAAACCGGCAAGAGACCAACATATATAGAGTACCGTGTGAGATTGAAT 2712
DB 558 GAGCAATGAGAGACACCATCTGACACATGACGTTGAGTGTGAGTCTGCAC 499
QY 2713 CAGTCCGTTGCCATTAACCAATCATATCTGCAATTGGGTGGAAGCA 2765
DB 498 CAAGACGTCACACACACCCGATCAGACCCCAACAAAGGTTGACATTA 446

RESULT 15
US-10-173-999-149/c
Sequence 149, Application US/10173999
Publication No. US2004005563A1
GENERAL INFORMATION:
APPLICANT: Mack, David H.
APPLICANT: Gish, Kurt C.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
FILE REFERENCE: 018501-002420US
CURRENT APPLICATION NUMBER: US/10/173,999
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: US 60/299,234
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/315,287
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/372,246
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 163
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 149
LENGTH: 2280
TYPE: DNA
ORGANISM: Homo sapiens
US-10-173-999-149

Query Match 2.6%; Score 75.4; DB 16; Length 2280;
Best Local Similarity 56.0%; Pred. No. 7.7e-13;
Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGTGAAGATGACCGGAGAAAGATTGAAAAATCGTGATGGTGC 2535
DB 771 CACCTCCAGGCGGACAGACCAACAGGACGCGAGTGAAGAGTCAATGACAGTGC 712
QY 2536 ACTGGGAAAGCAACGCGGAACTCTTTACAGCGCATATGCGCAAGGCTGACGAGT 2595
DB 721 TCCGTGAAAGCTCTTCTGTAATCTCCATCTCCACCTGCTGAGCGCAATAGT 652
QY 2596 ATTGCTCAGATGATCAATATTGGCAACCCATCAGCAATGAAATCGCGTTCAACCG- 2654
DB 651 GTTGGATGATGACCTTCCATATGTTGGCCCCCATATATGAGGATGACGCCGAAACGT 592
QY 2655 -GTAAACCGGCAAGAGACCAACATATATGAGTACCGTGTGAGATTGAAT 2712
DB 591 GAGCAATGAGAGACACCATCTGACACATGAGATGAGTGTGAGTCTGCAC 532
QY 2713 CAGTCCGTTGCCATTAACCAATCATATCTGCAATTGGGTGGAAGCA 2765
DB 531 CAAGACGTCACACACCCGATCAGACCCCAACAAAGGTTGACATTA 479

Search completed: July 31, 2004, 19:10:09
Job time : 903 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:32:25 ; Search time 761 Seconds
(without alignments)
16239.170 Million cell updates/sec

Title: US-09-700-712A-1

Sequence: 1 gagaaagtttctatgcctc.....caggaatcgcgctgcag 2909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2309	100.0	2909	AA240646	AA240646 Nucleotide
2	1222	42.0	1222	AA240648	AA240648 Nucleotide
3	1112.6	38.2	1149	AC653067	AC653067 Prokaryot
4	842.4	29.0	852	AA224801	AA224801 Vibrio ch
5	842.4	29.0	852	AC653066	AC653066 Prokaryot
6	838	28.8	838	AA240647	AA240647 Nucleotide
7	744.2	25.6	110000	AA240647	Continuation (10 o
8	505.4	17.4	852	AC653067	Continuation (10 o
9	462.2	15.9	852	AC653067	Continuation (10 o
10	403.8	13.9	849	AC653067	Continuation (10 o
11	292.2	10.0	807	AA237118	AA237118 Nucleotide
12	278.2	9.6	110000	AA237118	Continuation (28 o
13	278.2	9.6	110000	AA237118	Continuation (28 o
14	278.2	9.6	110000	AA237118	Continuation (28 o
15	277.8	9.5	876	AA237118	Continuation (2 of
16	248.2	8.5	876	AA237118	Continuation (2 of
17	208.6	7.2	834	AA237118	Continuation (2 of
18	208.4	7.2	834	AA237118	Continuation (2 of
19	161.2	5.5	812	AA237118	Continuation (2 of
20	156.4	5.4	852	AA237118	Continuation (2 of
21	156.4	5.4	102634	AA237118	Continuation (2 of
22	156.4	5.4	110000	AA237118	Continuation (11 o
23	156.4	5.4	343980	AA237118	Continuation (11 o

24	152.2	5.2	110000	6	AB292787_4	Continuation (5 of
25	149.2	5.1	849	7	AB240310	Ab240310 N. gonorr
26	149.2	5.1	852	3	AA253290	AA253290 Neisseria
27	148.6	5.1	1515	4	AA261079	AA261079 P. putida
28	145.4	5.0	420	6	AB26154	AB26154 Human ORF
29	142.2	4.9	96109	4	AA262548	AA262548 Human ORF
30	137	4.7	420	6	AB26154	AB26154 Human ORF
31	84	2.9	2440	9	AB241887	AB241887 Human ORF
32	84	2.9	2440	9	AB241887	AB241887 Human ORF
33	84	2.9	2440	9	AB241887	AB241887 Human ORF
34	83.6	2.9	4473	5	AA288385	AA288385 DNA encod
35	79	2.7	2075	9	AB288991	AB288991 Toxicity
36	79	2.7	2075	9	AB288991	AB288991 Toxicity
37	77.8	2.7	3950	7	AB242286	AB242286 Toxicity
38	75.4	2.6	2208	4	AA209555	AA209555 Human Na-
39	75.4	2.6	2208	4	AA209555	AA209555 Human Na-
40	75.4	2.6	2280	9	AB280585	AB280585 Ovarian c
41	75.4	2.6	2288	2	AA259498	AA259498 Human sod
42	75.4	2.6	2380	4	AA251764	AA251764 Human pol
43	75.4	2.6	2520	4	AA251764	AA251764 Human pol
44	75.4	2.6	3952	8	ACH03831	ACH03831 Human cDN
45	75.4	2.6	4137	3	AA275484	AA275484 DNA encod

ALIGNMENTS

RESULT 1
AA240646
ID AA240646 standard; DNA; 2909 BP.
XX
AC AA240646;
XX
DT 08-MAR-2000 (first entry)
XX
DE Nucleotide sequence of V. cholerae *thyA* gene.
XX
KM Vibrio cholerae; *thyA*-negative strain; Deltathya; *thyA* gene; vaccine;
XX
KW cholera; antimicrobial therapy; ss.
XX
OS Vibrio cholerae.
XX
PN WO9961634-A1.
XX
PD 02-DEC-1999.
XX
PF 21-MAY-1999; 99WO-EP003509.
XX
PR 26-MAY-1998; 98SE-00001852.
XX
PA (SBLV-) SBL VACCIN AB.
XX
PI Carlin N, Lebens MR;
XX
DR WPI, 2000-062719/05.
XX
DR P-PSDB; AAY59126.
XX
PT New Vibrio cholerae strain defective in the *thyA* gene, for use in
XX
PS vaccines and for recombinant protein production.
XX
PS Claim 9, Fig 1, 42pp; English.
XX
The invention provides a method for producing a Vibrio cholerae *thyA*-
XX
negative strain which is Deltathya strain lacking *thyA* gene functions.
XX
The method comprises site-directed mutagenesis of the V. cholerae
XX
chromosome to delete and/or insert nucleotides at the *thyA* locus; The V.
XX
chromosome *thyA*-negative strains are used: for overproduction of
XX
recombinant proteins; and in vaccines to prevent or treat cholera (or
XX
other diseases, if engineered to express the appropriate proteins). The
XX
thyA gene is also useful for insertion of foreign genes, in a selective
XX
and site-specific manner, and the proteins expressed by the *thyA* gene or
XX
by its 5'-flanking region, are useful in research and as targets for
XX
antimicrobial therapy. When used for recombinant protein production, V.

CC cholerae provides high yields with secretion of products into the culture medium for ease of subsequent recovery. The thA-negative strain can be maintained by thymine complementation, eliminating the need for antibiotic selection. The present sequence represents the nucleotide sequence of the V. cholerae thA gene

XX Sequence 2909 BP; 734 A; 644 C; 737 G; 794 T; 0 U; 0 Other;

Query Match 100.0%; Score 2909; DB 3; Length 2909;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GAGAGGTTGTTATGCTCAGAGGTTATCTGACATTTCCCAATTATGACCCCGATTTGTT 60
DB 1 GAGAGGTTGTTATGCTCAGAGGTTATCTGACATTTCCCAATTATGACCCCGATTTGTT 60
QY 61 TTGATCGGCGCTTACGCGGTGCGCTGTATGCTTATGTTATTTGGTGGTTTCTTTT 120
DB 61 TTGATCGGCGCTTACGCGGTGCGCTGTATGCTTATGTTATTTGGTGGTTTCTTTT 120
QY 121 TGTATGTTGGTGGCCAAATCGCGAGCGGATCGCGGGCAGTGTGTAAGCGGTGAGCA 180
DB 121 TGTATGTTGGTGGCCAAATCGCGAGCGGATCGCGGGCAGTGTGTAAGCGGTGAGCA 180
QY 181 AGTCTCTGACTTGTATTTGCGCGCTTTTAAAGTATGATCGGTGCGCAATTTGTTA 240
DB 181 AGTCTCTGACTTGTATTTGCGCGCTTTTAAAGTATGATCGGTGCGCAATTTGTTA 240
QY 241 TGTGATCTTCAATTTGATCTGTTCTGTGTAACCTTTATTTATTAAGATGTG 300
DB 241 TGTGATCTTCAATTTGATCTGTTCTGTGTAACCTTTATTTATTAAGATGTG 300
QY 241 TGTGATCTTCAATTTGATCTGTTCTGTGTAACCTTTATTTATTAAGATGTG 300
DB 301 GACTGCGGCGATGCTCTTCCAGCGCGCTTATTTGGGTGTGATCACCGCGCATTTGCTA 360
DB 301 GACTGCGGCGATGCTCTTCCAGCGCGCTTATTTGGGTGTGATCACCGCGCATTTGCTA 360
QY 361 TGGCGCTTAAACCAACGACCTCTTTTGTGTGCGCGCATTTGTTGCGCCCTTTAGTCC 420
DB 361 TGGCGCTTAAACCAACGACCTCTTTTGTGTGCGCGCATTTGTTGCGCCCTTTAGTCC 420
QY 421 TGGCGCTTAAACCAACGACCTCTTTTGTGTGCGCGCATTTGTTGCGCCCTTTAGTCC 420
DB 421 TGGCGCTTAAACCAACGACCTCTTTTGTGTGCGCGCATTTGTTGCGCCCTTTAGTCC 420
QY 421 ATTGCGTTGGGATGGAGCTATCGGTAACTTATGATAGTGAATTTGGGAGCGAGT 480
DB 421 ATTGCGTTGGGATGGAGCTATCGGTAACTTATGATAGTGAATTTGGGAGCGAGT 480
QY 481 AACGATGTGCTTGGGCTTTTGTATTTCCCTAATGATGCGCCACATGCGCGCATCTTC 540
DB 481 AACGATGTGCTTGGGCTTTTGTATTTCCCTAATGATGCGCCACATGCGCGCATCTTC 540
QY 541 AACGATGTGCTTGGGCTTTTGTATTTCCCTAATGATGCGCCACATGCGCGCATCTTC 540
DB 541 AACGATGTGCTTGGGCTTTTGTATTTCCCTAATGATGCGCCACATGCGCGCATCTTC 540
QY 541 ACAGCTTATGAAATGCGCTTGAAGCGGTGCTTCTTTCTTTATTTCTTAATTTGGTTAT 600
DB 541 ACAGCTTATGAAATGCGCTTGAAGCGGTGCTTCTTTCTTTATTTCTTAATTTGGTTAT 600
QY 601 TGGTAAACCTGTCGCGTAGCGAGCTATCGGACTGTTTTTAACTGATACGCTAATT 660
DB 601 TGGTAAACCTGTCGCGTAGCGAGCTATCGGACTGTTTTTAACTGATACGCTAATT 660
QY 661 CCGCTTCTTGTGAATAGTCCGTAGCCAGATGCTAGTGGGTCTTTTGGTGGCT 720
DB 661 CCGCTTCTTGTGAATAGTCCGTAGCCAGATGCTAGTGGGTCTTTTGGTGGCT 720
QY 661 CCGCTTCTTGTGAATAGTCCGTAGCCAGATGCTAGTGGGTCTTTTGGTGGCT 720
DB 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATCATGATATTTGATGATGAT 780
QY 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATCATGATATTTGATGATGAT 780
DB 781 TTGGCTTCAAGCGCGGTTTGTATCAAGCGGTGAGCAAAATAGGATAGT 840
QY 781 TTGGCTTCAAGCGCGGTTTGTATCAAGCGGTGAGCAAAATAGGATAGT 840
DB 841 GAAACGATATTTAGATTTTGTACAGCGCATCTGATCAAGGCTTTGGTTGAAATGA 900
QY 841 GAAACGATATTTAGATTTTGTACAGCGCATCTGATCAAGGCTTTGGTTGAAATGA 900
DB 901 ACAGACGGGCAAGCGTTTGACTGTATTAATGCCGATTTGACCTAGCATGTGGCA 960
```

```
DB 901 ACAGACGGGCAAGCGTTTGACTGTATTAATGCCGATTTGACCTAGCATGTGGCA 960
QY 961 CATATGTTTCCCTTAGATGATCAACGAGAGTTTGTGAAGCTGCGGTAGCGAGT 1020
DB 961 CATATGTTTCCCTTAGATGATCAACGAGAGTTTGTGAAGCTGCGGTAGCGAGT 1020
QY 961 CATATGTTTCCCTTAGATGATCAACGAGAGTTTGTGAAGCTGCGGTAGCGAGT 1020
DB 1021 GCTCGGCTAATTCGTGTTTGAATATGCGGCGGATTTTCCCAATTAGTATCCAAAC 1080
QY 1021 GCTCGGCTAATTCGTGTTTGAATATGCGGCGGATTTTCCCAATTAGTATCCAAAC 1080
DB 1081 CTGGATGCTAATGCAATTTAAACCAACATGCTGCAACATCTTACCTGATAGGTTA 1140
QY 1081 CTGGATGCTAATGCAATTTAAACCAACATGCTGCAACATCTTACCTGATAGGTTA 1140
DB 1141 GATGACATGAGACGCGGTATGTTGATGATTTGATGATTTGATGATTTGATGATTTG 1200
QY 1141 GATGACATGAGACGCGGTATGTTGATGATTTGATGATTTGATGATTTGATGATTTG 1200
DB 1201 TCAATATTGACCACTTGAATAAGATTTGATGATTTGATGATTTGATGATTTGATGAT 1260
QY 1201 TCAATATTGACCACTTGAATAAGATTTGATGATTTGATGATTTGATGATTTGATGAT 1260
DB 1261 TGAATTTCTTACCTTACATCGGCTGATTTTCAATGAGGTGTTTGGCGCTTGAT 1320
QY 1261 TGAATTTCTTACCTTACATCGGCTGATTTTCAATGAGGTGTTTGGCGCTTGAT 1320
DB 1261 TGAATTTCTTACCTTACATCGGCTGATTTTCAATGAGGTGTTTGGCGCTTGAT 1320
QY 1321 GATACGCAATCATTTTCAATGCTGCGGATATCTTGTATCTTCAACAGATCTGATG 1380
DB 1321 GATACGCAATCATTTTCAATGCTGCGGATATCTTGTATCTTCAACAGATCTGATG 1380
QY 1381 ATGTGATGTCCTTGGGCTTGAATTTCAACATGATGATGATTTGATGATTTGATGAT 1440
DB 1381 ATGTGATGTCCTTGGGCTTGAATTTCAACATGATGATGATTTGATGATTTGATGAT 1440
QY 1441 GATGCAACATCAACGAGGAAAGCGGCTTGGCGATATCAACAGATCTGATGATG 1500
DB 1441 GATGCAACATCAACGAGGAAAGCGGCTTGGCGATATCAACAGATCTGATGATG 1500
QY 1501 CATTTACCAATCACTGATTTGATGCGGATATGCGGATATGCGGATATGCGGATATG 1560
DB 1501 CATTTACCAATCACTGATTTGATGCGGATATGCGGATATGCGGATATGCGGATATG 1560
QY 1561 AGGCTCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1620
DB 1561 AGGCTCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1620
QY 1621 CACTTGAATGATTTGATGCTACCGGATATGCTTCAACAGATCTGATGATGATG 1680
DB 1621 CACTTGAATGATTTGATGCTACCGGATATGCTTCAACAGATCTGATGATGATG 1680
QY 1681 TTAGCTAATCCGATTTAGCGGCTGATGAGGTTTATTAATTAATTAATTAATTAAT 1740
DB 1681 TTAGCTAATCCGATTTAGCGGCTGATGAGGTTTATTAATTAATTAATTAATTAAT 1740
QY 1741 CGAAGTCTGGAGCTTTTATTAACAGATGCTTAAAGCTTTAGGCTGAG 1800
DB 1741 CGAAGTCTGGAGCTTTTATTAACAGATGCTTAAAGCTTTAGGCTGAG 1800
QY 1801 AATGCTCGGAGTGAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1860
DB 1801 AATGCTCGGAGTGAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1860
QY 1861 CTTAACAAGCCCAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920
DB 1861 CTTAACAAGCCCAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920
QY 1921 AATACCGTAAATCAAGCCGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
DB 1921 AATACCGTAAATCAAGCCGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
QY 1981 CAGAGCAACCGGCAATCTACAGAGACGCGGTTCCGCGAGCAAGACGATATGACA 2040
```

```

Db      1981 CAGAGCAACAGCGCAAACTCACCGAGACAGCGGTTGGCGGAGAGAGAGCTAATAC 2040
Qy      2041 AGTGGCAATTTGGCACTTAAGTAAATGGGTAGATTTCACGCACTTTGACAGCGCAGA 2100
Db      2041 AGTGGCAATTTGGCACTTAAGTAAATGGGTAGATTTCACGCACTTTGACAGCGCAGA 2100
Qy      2101 GCCCAGAGAGAGAACCATTAAGGCTGGTGTGGTGCATGAAGATTGAATTAATACGCTAAC 2160
Db      2101 GCCCAGAGAGAGAACCATTAAGGCTGGTGTGGTGCATGAAGATTGAATTAATACGCTAAC 2160
Qy      2161 CACTGTACTGTAAGCAATAACCGTGTAGTGGGCTCGGCCAATGCAATTTTGTAGAAATTC 2220
Db      2161 CACTGTACTGTAAGCAATAACCGTGTAGTGGGCTCGGCCAATGCAATTTTGTAGAAATTC 2220
Qy      2221 AGGTGGCGGGCCACCATCAAACTCTTCATCAGTTTGGTGGCCATACCGCTAATGGCGACGA 2280
Db      2221 AGGTGGCGGGCCACCATCAAACTCTTCATCAGTTTGGTGGCCATACCGCTAATGGCGACGA 2280
Qy      2281 AATGTGCGAATACCCCAATACGATAAGTGCACACCCAGCAAGATTAACCAATACGA 2340
Db      2281 AATGTGCGAATACCCCAATACGATAAGTGCACACCCAGCAAGATTAACCAATACGA 2340
Qy      2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGGTAAATCGGTTTGAATTAATCAAACC 2400
Db      2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGGTAAATCGGTTTGAATTAATCAAACC 2400
Qy      2401 TTTTCATGCTCATATCGCACTGCGCAAGAGCGGCAAGAGCGAGTGTGAATTTCTC 2460
Db      2401 TTTTCATGCTCATATCGCACTGCGCAAGAGCGGCAAGAGCGAGTGTGAATTTCTC 2460
Qy      2461 TAAATATCCAAATCATATTTCTAGAGGTAGAAAGATGACACCGCGAGAAATTGAAAA 2520
Db      2461 TAAATATCCAAATCATATTTCTAGAGGTAGAAAGATGACACCGCGAGAAATTGAAAA 2520
Qy      2521 ATCGTGGATGGTGGCACTGGCGCAAGAGCGGCAAGAGCGGCAAGAGCGGCAAGAGCGGCA 2580
Db      2521 ATCGTGGATGGTGGCACTGGCGCAAGAGCGGCAAGAGCGGCAAGAGCGGCAAGAGCGGCA 2580
Qy      2581 AAGGCTGACAGAGATTTGGTCAACAGTATACCAATATTGCAACCATCACCATAGGAAT 2640
Db      2581 AAGGCTGACAGAGATTTGGTCAACAGTATACCAATATTGCAACCATCACCATAGGAAT 2640
Qy      2641 CGCGGTTTAAACCGGTAACCCACCGGCAACGAGCAACCAATTAATGAAGTACCGTGCT 2700
Db      2641 CGCGGTTTAAACCGGTAACCCACCGGCAACGAGCAACCAATTAATGAAGTACCGTGCT 2700
Qy      2701 TGAAGATTGAATCAGTGGCGTGGCCACTAAACCAATCATCAATCTGCAATTTGGGTGGA 2760
Db      2701 TGAAGATTGAATCAGTGGCGTGGCCACTAAACCAATCATCAATCTGCAATTTGGGTGGA 2760
Qy      2761 AGCAATTTCAATTAAGACTTTGGCTTGAATGCGCGGTTGGCCATTTAAACCGCTCGAC 2820
Db      2761 AGCAATTTCAATTAAGACTTTGGCTTGAATGCGCGGTTGGCCATTTAAACCGCTCGAC 2820
Qy      2821 CATCGGACATGCAAGAGATAGTAATATACGATGAAAGCCAAATTTGGCCAAAGTAGGCC 2880
Db      2821 CATCGGACATGCAAGAGATAGTAATATACGATGAAAGCCAAATTTGGCCAAAGTAGGCC 2880
Qy      2881 TTTTCGTGTCAGCGAAATCGCGCTGACG 2909
Db      2881 TTTTCGTGTCAGCGAAATCGCGCTGACG 2909

```

```

RESULT 2
AAZ40648
ID      AAZ40648 standard; DNA; 1222 BP.
XX
AC      AAZ40648;
XX
DT      08-MAR-2000 (first entry)
XX
DE      Nucleotide sequence of V. cholerae thya gene 3' flanking region.
XX

```

```

KM      Vibrio cholerae; thya-negative strain; Deltathya; thya gene; vaccine;
KM      cholera; antimicrobial therapy; ss.
XX
OS      Vibrio cholerae.
XX
PN      MO9961634-A1.
XX
PD      02-DEC-1999.
XX
PF      21-MAY-1999; 99WO-EP003509.
XX
PR      26-MAY-1998; 98SE-00001852.
XX
PA      (SBLV-) SBL VACCIN AB.
XX
PI      Carlin N, Lebens MR;
XX
DR      WPI; 2000-062719/05.
XX
PT      New Vibrio cholerae strain defective in the thya gene, for use in
PT      vaccines and for recombinant protein production.
XX
PS      Claim 11; Fig 3; 42pp; English.
XX
CC      The invention provides a method for producing a Vibrio cholerae thya-
CC      negative strain which is Deltathya strain lacking thya gene functions.
CC      The method comprises site-directed mutagenesis of the V. cholerae
CC      chromosome to delete and/or insert nucleotides at the thya locus; The V.
CC      cholerae thya-negative strains are used: for overproduction of
CC      recombinant proteins; and in vaccines to prevent or treat cholera (or
CC      other diseases if engineered to express the appropriate proteins). The
CC      thya gene is also useful for insertion of foreign genes, in a selective
CC      and site-specific manner, and the proteins expressed by the thya gene or
CC      by its 5'-flanking region, are useful in research and as targets for
CC      antimicrobial therapy. When used for recombinant protein production, V.
CC      cholerae provides high yields with secretion of products into the culture
CC      medium for ease of subsequent recovery. The thya-negative strain can be
CC      maintained by thymine complementation, eliminating the need for
CC      antibiotic selection. The present sequence represents the nucleotide
CC      sequence of the V. cholerae thya gene 3' flanking region
XX
SQ      Sequence 1222 BP; 373 A; 293 C; 289 G; 267 T; 0 U; 0 Other;
XX
Query Match      42.0%; Score 1222; DB 3; Length 1222;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1688 TAAATCCGTAATCAAGGCGGTATGCTTGAATGAGGTTTATTAATTAAGTCCCGAAGT 1747
Db      1 TAAATCCGTAATCAAGGCGGTATGCTTGAATGAGGTTTATTAATTAAGTCCCGAAGT 60
Qy      1748 CGGAGCTTTTATTAATCAAGATGATGCTTAACGCTTAAGCGTTAGGGCAAGATCTG 1807
Db      61 CGGAGCTTTTATTAATCAAGATGATGCTTAACGCTTAAGCGTTAGGGCAAGATCTG 120
Qy      1808 CGGAGGATGACGACAAACACCAATTAAGTCAACCAACCATTTTGTCTTACAA 1867
Db      121 CGGAGGATGACGACAAACACCAATTAAGTCAACCAACCATTTTGTCTTACAA 180
Qy      1868 GCCCAAGTTAGATGAGTACGACACCTTAATAGGCGAGTGGGTGAAGAAAGAAATACCG 1927
Db      181 GCCCAAGTTAGATGAGTACGACACCTTAATAGGCGAGTGGGTGAAGAAAGAAATACCG 240
Qy      1928 TAAATCAAGACCGTAGGCATCAAGTTAAAGCTTAAGTGACACGACCAATTTGACAGCA 1987
Db      241 TAAATCAAGACCGTAGGCATCAAGTTAAAGCTTAAGTGACACGACCAATTTGACAGCA 300
Qy      1988 AACACGGCAAACTCAACAGAGACAGCGGTTGGCGGACGACGACGATTAATACAGTGCCA 2047
Db      301 AACACGGCAAACTCAACAGAGACAGCGGTTGGCGGACGACGACGATTAATACAGTGCCA 360
Qy      2048 ATGTTCGACACTAAGTAATGGTATGATTTCACGCACTTTGACGACGCCAGGCCACG 2107

```



```
QY 1903 CAGTTGGGTAAAGAAAGATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAA 1962
DB 1029 CAGTTGGGTAAAGAAAGATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAA 970
QY 1963 GTGCAACGACGCAATTTGGAGAGCAAAACGCGCAAACTCCAGAGACAGCGGTTGCGGC 2022
DB 969 GTGCAACGACGCAATTTGGAGAGCAAAACGCGCAAACTCCAGAGACAGCGGTTGCGGC 910
QY 2023 GAGCAGAGAGTAATACAAAGTCCCAATGTTCCACCTAAGTAATGGTGTGATTTTCAG 2082
DB 909 GAGCAGAGAGTAATACAAAGTCCCAATGTTCCACCTAAGTAATGGTGTGATTTTCAG 850
QY 2083 CACTTTCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2142
DB 849 CACTTTCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 790
QY 2143 TTGAACTAATACCGTAACCTGTACCTGAGCAATACCGTGTAGTGGGCTCGGCCAT 2202
DB 789 TTGAACTAATACCGTAACCTGTACCTGAGCAATACCGTGTAGTGGGCTCGGCCAT 730
QY 2203 CGCATTTTGTAGAAATTTTCAGTGGGCGGCAACCACTCAACTCTTCATCAAGTTGCCAT 2262
DB 729 CGCATTTTGTAGAAATTTTCAGTGGGCGGCAACCACTCAACTCTTCATCAAGTTGCCAT 670
QY 2263 CACCGTAATGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2322
DB 669 CACCGTAATGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 610
QY 2322 AGTATTAACCAATACCGAAAGTGGGTTTCAGACCCCTGAGAGAGAGTTGGTAATGG 2382
DB 609 AGTATTAACCAATACCGAAAGTGGGTTTCAGACCCCTGAGAGAGAGTTGGTAATGG 550
QY 2383 TTTGATTAATCAAAACCTTTCATGCTCATATGCGCATGCGACAGAGCGGCAAGCAG 2442
DB 549 TTTGATTAATCAAAACCTTTCATGCTCATATGCGCATGCGACAGAGCGGCAAGCAG 490
QY 2443 CCAAGTGTAGACTTTCTTAAATATGCAACATCATTTTAAAGGTAGAGATAGACAC 2502
DB 489 CCAAGTGTAGACTTTCTTAAATATGCAACATCATTTTAAAGGTAGAGATAGACAC 430
QY 2503 CCGGAGAAATGAGAAATATGAGATGAGTGGCACTGGGCAAGAGACGCGCAAACTTTC 2562
DB 429 CCGGAGAAATGAGAAATATGAGATGAGTGGCACTGGGCAAGAGACGCGCAAACTTTC 370
QY 2563 TTTAAGCGCATATGCGCAAGGCTGACGAGATTTGTCACATGATCCAAATTTGAC 2622
DB 369 TTTAAGCGCATATGCGCAAGGCTGACGAGATTTGTCACATGATCCAAATTTGAC 310
QY 2623 ACCCATCAACATAGGAATGCGGTTTCACCGGTAACCCAGCGCAAGAGCAACAT 2682
DB 309 ACCCATCAACATAGGAATGCGGTTTCACCGGTAACCCAGCGCAAGAGCAACAT 250
QY 2683 AATAGAGTACCGTCTTGAAGATTTGAATCAGTGGCGGTTGCACTAAACCAATCA 2742
DB 249 AATAGAGTACCGTCTTGAAGATTTGAATCAGTGGCGGTTGCACTAAACCAATCA 190
QY 2743 TCTGTGAATTTGGGTGGGAAACCAATTTCAATTTGAATTTGGGCTTGAATGCGGCTGCCA 2802
DB 189 TCTGTGAATTTGGGTGGGAAACCAATTTCAATTTGAATTTGGGCTTGAATGCGGCTGCCA 130
QY 2803 TTTAAACCGCTGCGACCATGCGCACTGCAAGAGTAAATACAGATGAAAGCCAA 2862
DB 129 TTTAAACCGCTGCGACCATGCGCACTGCAAGAGTAAATACAGATGAAAGCCAA 70
QY 2863 GTTTGCGCAAGTGAAGGCTTTGTGTGTCAGCGAATCGGCGCTGCA 2909
DB 69 GTTTGCGCAAGTGAAGGCTTTGTGTGTCAGCGAATCGGCGCTGCA 23
```

RESULT 4
AAZ24801
ID AAZ24801 standard, DNA; 852 BP.
XX

```
AC AAZ24801;
XX
DT 07-DEC-1999 (first entry)
XX
DE Vibrio cholerae thya coding region sequence.
XX
KW Wild type; thya; enzyme; thymidilate synthase; non-toxicogenic; mutation;
KW attenuation; immunization; cholera; hemagglutinin protease; biosafety;
KW auxotrophism; replication; vaccine; ss.
XX
OS Vibrio cholerae.
XX
PN MO9935271-A2.
XX
PD 15-JUL-1999.
XX
PE 30-DEC-1998; 98MO-CU000008.
XX
PR 30-DEC-1997; 97CU-00000142.
XX
PA (NAIN-) CENT NACIONAL INVESTIGACIONES CIENTIFICA.
XX
PI Campos Gomez J, Fando Galzada RA, Rodriguez Gonzalez BJ;
PI Ledon Perez TY, Valle Diaz E, Silva Cabrera AJ, Benitez Robles JA;
DR WP1; 1999-430398/36.
XX
DR P-PSDB; AAY26895.
XX
PT Producing strains of Vibrio cholerae with inactivated gene for
PT hemagglutinin protease, useful in vaccines against cholera.
XX
PS Claim 18; Page 30; 30pp; Spanish.
XX
SS
XX
CC This sequence represents the wild type coding region of the thya gene
CC from Vibrio cholerae, which encodes the enzyme thymidilate synthase. The
CC invention relates to the production, from a non-toxicogenic strain of V.
CC cholerae, of attenuated strains of V. cholerae suitable for immunization
CC against cholera. The attenuation comprises inactivating the gene for
CC hemagglutinin protease (HP), either by deletion, insertion or some other
CC defined and irreversible genetic manipulation. Additional biosafety of
CC the attenuated strains is generated by mutating the thya gene, resulting
CC in auxotrophic mutants unable to replicate in the environment. The new
CC strains are used to produce anticholera vaccines
XX
SQ Sequence 852 BP; 217 A; 177 C; 216 G; 242 T; 0 U; 0 Other;
Query Match 29.0%; Score 842.4; DB 2; Length 852;
Best Local Similarity 99.3%; Pred. No. 1.7e-257;
Matches 846; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 839 GTGAAACAGATTTTAATCTTTGTACGCGCATGTCATCAAGGTGTTGGTTGAAAT 898
DB 1 GTGAAACAGATTTTAATCTTTGTACGCGCATGTCATCAAGGTGTTGGTTGAAAT 60
QY 899 GAAAGAACGCGCAAGCGTTGTTGACTGTGATTAATCCGATTGACATGTCGCG 958
DB 61 GAAAGAACGCGCAAGCGTTGTTGACTGTGATTAATCCGATTGACATGTCGCG 120
QY 959 AACAAATAGTTTCTCTAGTACTACGCAAGATTTTGGAAAGTGCCTGAGCCGAG 1018
DB 121 AACAAATAGTTTCTCTAGTACTACGCAAGATTTTGGAAAGTGCCTGAGCCGAG 180
QY 1019 TTGCTGGCTAATTTGCTGTGTTAGTAATGCGGCGGATTTTGGCAATTAGGACCAA 1078
DB 181 TTGCTGGCTAATTTGCTGTGTTAGTAATGCGGCGGATTTTGGCAATTAGGACCAA 240
QY 1079 ACCTGGATGCTAATGCAATTTAAACCAAGCATGCTCAACAATCTTAACGTTAAAGT 1138
DB 241 ACCTGGATGCTAATGCAATTTAAACCAAGCATGCTCAACAATCTTAACGTTAAAGT 300
QY 1139 GAGGATGACATGGGACGCGGTATGAGTCAAGGTAGAGCTTGGGCTAAGCCTGATGCT 1198
DB 301 GAGGATGACATGGGACGCGGTATGAGTCAAGGTAGAGCTTGGGCTAAGCCTGATGCT 360
```

QY 1199 GGTGATATGACCAAGTGAAGAAAGATTGTTGATGATTTGAGCCGTGGCCTTGATGACCGA 1258
 DB 361 GGTGATATGACCAAGTGAAGAAAGATTGTTGATGATTTGAGCCGTGGCCTTGATGACCGA 420
 QY 1259 GGTGAATCTTACTCTTCAACATCCGGGTGAATTTCAATGGGGTGGTGGCCCTTGC 1318
 DB 421 GGTGAATCTTACTCTTCAACATCCGGGTGAATTTCAATGGGGTGGTGGCCCTTGC 480
 QY 1319 ATGTACAGCCATCATTTTTCATTTGCTGGGGGATACCTTGTATCTCAACAGTACGAGCT 1378
 DB 481 ATGTACAGCCATCATTTTTCATTTGCTGGGGGATACCTTGTATCTCAACAGTACGAGCT 540
 QY 1379 TCATGTGATGTCCTTGGGGTGGTGAATTTTCAACATGGTGCAGTTTATGTTCTTGGC 1438
 DB 541 TCATGTGATGTCCTTGGGGTGGTGAATTTTCAACATGGTGCAGTTTATGTTCTTGGC 600
 QY 1439 CTGATGCGCAGATTCACAGGGAAGAAAGCCGGCTTGGCTATCAACAGATCTGATCGC 1498
 DB 601 CTGATGCGCAGATTCACAGGGAAGAAAGCCGGCTTGGCTATCAACAGATCTGATCGC 660
 QY 1499 CACATTTACCAAGATTAATCTCGATGATGCGGATGTCAGCTTAAACGTGAGCCATTC 1558
 DB 661 CACATTTACCAAGATTAATCTCGATGATGCGGATGTCAGCTTAAACGTGAGCCATTC 720
 QY 1559 CCAGGCGCTCAGTTCATATCAATCCAAAGATTAAACACTGACGATTTGGAACCTTGC 1618
 DB 721 CCAGGCGCTCAGTTCATATCAATCCAAAGATTAAACACTGACGATTTGGAACCTTGC 780
 QY 1619 GTGCTTTGGATGATTTTGAACGTACCGGATTAACGATTCACGATCTTAAATACCGC 1678
 DB 781 GTGCTTTGGATGATTTTGAACGTACCGGATTAACGATTCACGATCTTAAATACCGC 840
 QY 1679 TTTTCAGTCTAA 1690
 DB 841 TTTTCAGTCTAA 852

RESULT 5

ACAS3066
 ID ACAS3066 standard; DNA; 852 BP.
 ACAS3066;
 19-JUN-2003 (first entry)
 DE Prokaryotic essential gene #34723.
 KM Antisense; ds; prokaryotic essential gene; cell proliferation;
 KM drug design; gene.
 OS Vibrio cholerae.
 PN W020027783-A2.
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002WO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0142923P.
 PR 06-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
 PI Wall D, Treweek JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU49196.
 XX

PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids, required for cellular proliferation or
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 14; SEQ ID NO 40936; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 613 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/publishes_pcr_sequences

XX Sequence 852 BP; 217 A; 177 C; 216 G; 242 T; 0 U; 0 Other;

XX Query Match 29.0%; Score 842.4; DB 7; Length 852;

XX Best Local Similarity 99.3%; Pred. No. 1.7e-257;

XX Matches 846; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 839 GTGAACAGTATTAGATCTTTGTCAAGCCGATGTGATCAAGGTTGGGTGAAGAT 898
 DB 1 GTGAACAGTATTAGATCTTTGTCAAGCCGATGTGATCAAGGTTGGGTGAAGAT 60
 QY 899 GAACGAACGGGCAAGGCTTGTGACGTGATTAATGCCGATTTGGAAGCTGCCGAG 958
 DB 61 GAACGAACGGGCAAGGCTTGTGACGTGATTAATGCCGATTTGGAAGCTGCCGAG 120
 QY 959 AACATCAGTTTCTCTAGTACTACACGCAAGATTTTGAAGCTGCCGAG 1018
 DB 121 AACATCAGTTTCTCTAGTACTACACGCAAGATTTTGAAGCTGCCGAG 180
 QY 1019 TTGCTGGCTAATTTGCTGTTAGATTAATGCCGAGATTTTGAAGCTGCCGAG 1078
 DB 181 TTGCTGGCTAATTTGCTGTTAGATTAATGCCGAGATTTTGAAGCTGCCGAG 240
 QY 1079 ACCTGGATCTAATGCCAATTTAAACCAAGCATGGCTCAACATCTTACCGTAAAGT 1138
 DB 241 ACCTGGATCTAATGCCAATTTAAACCAAGCATGGCTCAACATCTTACCGTAAAGT 300
 QY 1139 GAGGATACATGGGACCGCTGTATGTTGTTCAAGGTTAGAGCTTGGGCTAAGCTGATGT 1198
 DB 301 GAGGATACATGGGACCGCTGTATGTTGTTCAAGGTTAGAGCTTGGGCTAAGCTGATGT 360
 QY 1199 GGTGATATGACCAAGTGAAGAAAGATTGTTGATGATTTGAGCCGTGGCCTTGATGACCGA 1258
 DB 361 GGTGATATGACCAAGTGAAGAAAGATTGTTGATGATTTGAGCCGTGGCCTTGATGACCGA 420
 QY 1259 GGTGAATCTTACTCTTCAACATCCGGGTGAATTTCAATGGGGTGGTGGCCCTTGC 1318

Db 421 GGTGAATTTCTTAACCTTCAACATCCGGGTGAATTCACATGGGGTGTGGCCCTTGC 480
 QY 1319 ATGTACAGCCATCATTTTTCATTGCTGGGGGATACCTTGTATCTCAACAGTACTGAGCGT 1378
 Db 481 ATGTACAGCCATCATTTTTCATTGCTGGGGTATACCTTGTATCTCAACAGTACTGAGCGT 540
 QY 1379 TATATGTATGTGCTGGGGTGAATTTCAACATGTGCGAGTTATGTGTCCTTCCG 1438
 Db 541 TCATGTATGTGCTGGGGTGAATTTCAACATGTGCGAGTTATGTGTCCTTCCG 600
 QY 1439 CTGATGCGACAGATCAAGGAAAGCGGGCTTGCGTATCAACAGATGTCATGCG 1498
 Db 601 CTGATGCGACAGATCAAGGAAAGCGGGCTTGCGTATCAACAGATGTCATGCG 660
 QY 1499 CACATTTACCAAGATCACTGATGATGCGCGATGTGCGACTTAAACGTGAGCATTC 1558
 Db 661 CACATTTACCAAGATCACTGATGATGCGCGATGTGCGACTTAAACGTGAGCATTC 720
 QY 1559 CAGAGGCGCTCAGTCCATATCATCCAAAGATTAAACATGCGAGATTTGGAACTTGG 1618
 Db 721 CAGAGGCGCTCAGTCCATATCATCCAAAGATTAAACATGCGAGATTTGGAACTTGG 780
 QY 1619 GTCACTTTGATGATTTTTCAGTCAACCGGATATCAGTTCCAGATCTTATCAATACCG 1678
 Db 781 GTCACTTTGATGATTTTTCAGTCAACCGGATATCAGTTCCAGATCTTATCAATACCG 840
 QY 1679 TTTTCAGTCTAA 1690
 Db 841 TTTTCAGTCTAA 852
 RESULT 6
 AA240647
 ID AA240647 standard; DNA; 838 BP.
 AC AA240647;
 XX
 DT 08-MAR-2000 (first entry)
 XX
 DE Nucleotide sequence of V. cholerae thya gene 5' flanking region.
 XX
 KM Vibrio cholerae; thya-negative strain; Deltathya; thya gene; vaccine;
 KM cholera; antimicrobial therapy; ss.
 XX
 OS Vibrio cholerae.
 XX
 PN MO961634-AI.
 XX
 XX 02-DEC-1999.
 PD
 XX 21-MAY-1999; 99WO-EP003509.
 PF
 XX 26-MAY-1998; 98SE-00001852.
 PR
 XX (SBLV-) SBL VACCIN AB.
 PA
 XX Carlin N, Lebens MR;
 XX
 XX WPI; 2000-062719/05.
 DR P-PSDB; AAY59127.
 XX
 PT New Vibrio cholerae strain defective in the thya gene, for use in
 PT vaccines and for recombinant protein production.
 XX
 PS Claim 10; Fig 2; 42pp; English.
 CC The invention provides a method for producing a Vibrio cholerae thya-
 CC negative strain which is Deltathya strain lacking thya gene functions.
 CC The method comprises site-directed mutagenesis of the V. cholerae
 CC chromosome to delete and/or insert nucleotides at the thya locus; the V.
 CC cholerae thya-negative strains are used; for overproduction of
 CC recombinant proteins; and in vaccines to prevent or treat cholera (or

CC other diseases if engineered to express the appropriate proteins). The
 CC thya gene is also useful for insertion of foreign genes, in a selective
 CC and site-specific manner, and the proteins expressed by the thya gene or
 CC by its 5'-flanking region, are useful in research and as targets for
 CC antimicrobial therapy. When used for recombinant protein production, V.
 CC cholerae provides high yields with secretion of products into the culture
 CC medium for ease of subsequent recovery. The thya-negative strain can be
 CC maintained by thymine complementation, eliminating the need for
 CC antibiotic selection. The present sequence represents the nucleotide
 CC sequence of the V. cholerae thya gene 5' flanking region
 XX
 SQ Sequence 838 BP; 145 A; 176 C; 232 G; 285 T; 0 U; 0 Other;
 Query Match 28.8%; Score 838; DB 3; Length 838;
 Best Local Similarity 100.0%; Pred. No. 4,4e-256;
 Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAAAGTTTGTATGCTTCAAGGATATCTGCAAGTTTCCCAATATGACCCCGATTTGTT 60
 Db 1 GAGAAAGTTTGTATGCTTCAAGGATATCTGCAAGTTTCCCAATATGACCCCGATTTGTT 60
 QY 61 TTGATGCGGCGCTCTAGCGGTGCGTGTGATGCTGATGATTTGGTGGTTCCTTT 120
 Db 61 TTGATGCGGCGCTCTAGCGGTGCGTGTGATGCTGATGATTTGGTGGTTCCTTT 120
 QY 61 TTGATGCGGCGCTCTAGCGGTGCGTGTGATGCTGATGATTTGGTGGTTCCTTT 120
 Db 61 TTGATGCGGCGCTCTAGCGGTGCGTGTGATGCTGATGATTTGGTGGTTCCTTT 120
 QY 121 TGCTATGTGTGCGCAATGCGGAGCGATGCGCGGAGTGGTTGACGCGGTGAGCA 180
 Db 121 TGCTATGTGTGCGCAATGCGGAGCGATGCGCGGAGTGGTTGACGCGGTGAGCA 180
 QY 121 TGCTATGTGTGCGCAATGCGGAGCGATGCGCGGAGTGGTTGACGCGGTGAGCA 180
 Db 121 TGCTATGTGTGCGCAATGCGGAGCGATGCGCGGAGTGGTTGACGCGGTGAGCA 180
 QY 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGCGGCGGATGCTTGA 240
 Db 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGCGGCGGATGCTTGA 240
 QY 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGCGGCGGATGCTTGA 240
 Db 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGCGGCGGATGCTTGA 240
 QY 241 TGTGATCTTCAATTTTGTATGCTTCTGTCGACCTCTTATTTATTTCAAGTGTG 300
 Db 241 TGTGATCTTCAATTTTGTATGCTTCTGTCGACCTCTTATTTATTTCAAGTGTG 300
 QY 241 TGTGATCTTCAATTTTGTATGCTTCTGTCGACCTCTTATTTATTTCAAGTGTG 300
 Db 241 TGTGATCTTCAATTTTGTATGCTTCTGTCGACCTCTTATTTATTTCAAGTGTG 300
 QY 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGTGTATCACCGCCATGTTCTGTA 360
 Db 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGTGTATCACCGCCATGTTCTGTA 360
 QY 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGTGTATCACCGCCATGTTCTGTA 360
 Db 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGTGTATCACCGCCATGTTCTGTA 360
 QY 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGTGCGGATTTGTCCTTATGAGGCC 420
 Db 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGTGCGGATTTGTCCTTATGAGGCC 420
 QY 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGTGCGGATTTGTCCTTATGAGGCC 420
 Db 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGTGCGGATTTGTCCTTATGAGGCC 420
 QY 421 ATTGCGTTTGGGAGTGGAGCGTATCGGTAACTTTATGATATGAACTTTGGGAGCAGT 480
 Db 421 ATTGCGTTTGGGAGTGGAGCGTATCGGTAACTTTATGATATGAACTTTGGGAGCAGT 480
 QY 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGGCCACTGCGCGGCAATCCTTC 540
 Db 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGGCCACTGCGCGGCAATCCTTC 540
 QY 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGGCCACTGCGCGGCAATCCTTC 540
 Db 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGGCCACTGCGCGGCAATCCTTC 540
 QY 541 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGGCCACTGCGCGGCAATCCTTC 600
 Db 541 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGGCCACTGCGCGGCAATCCTTC 600
 QY 601 TGTAAACCTTGTGCTTGGGCTTGGAGCGGATCGGACCTTTTATGCTGATATGCTTAT 660
 Db 601 TGTAAACCTTGTGCTTGGGCTTGGAGCGGATCGGACCTTTTATGCTGATATGCTTAT 660
 QY 601 TGTAAACCTTGTGCTTGGGCTTGGAGCGGATCGGACCTTTTATGCTGATATGCTTAT 660
 Db 601 TGTAAACCTTGTGCTTGGGCTTGGAGCGGATCGGACCTTTTATGCTGATATGCTTAT 660
 QY 661 CCGCTTCTTGTGAATACGTCCTGAGCCGAGAGCTCAAGTGGTGTGTTGGTGGCTT 720
 Db 661 CCGCTTCTTGTGAATACGTCCTGAGCCGAGAGCTCAAGTGGTGTGTTGGTGGCTT 720
 QY 661 CCGCTTCTTGTGAATACGTCCTGAGCCGAGAGCTCAAGTGGTGTGTTGGTGGCTT 720
 Db 661 CCGCTTCTTGTGAATACGTCCTGAGCCGAGAGCTCAAGTGGTGTGTTGGTGGCTT 720
 QY 721 CATTTCAATGAGGCAATCTCTCTTACCTATGATATCGATATTTTGAATGATGCT 780
 Db 721 CATTTCAATGAGGCAATCTCTCTTACCTATGATATCGATATTTTGAATGATGCT 780
 QY 721 CATTTCAATGAGGCAATCTCTCTTACCTATGATATCGATATTTTGAATGATGCT 780
 Db 721 CATTTCAATGAGGCAATCTCTCTTACCTATGATATCGATATTTTGAATGATGCT 780
 QY 781 TTGCTTACAGAGCGGCTTGTATCAAGACCGTGTGACGCAAAATAGGTAAGTATG 838
 Db 781 TTGCTTACAGAGCGGCTTGTATCAAGACCGTGTGACGCAAAATAGGTAAGTATG 838
 QY 781 TTGCTTACAGAGCGGCTTGTATCAAGACCGTGTGACGCAAAATAGGTAAGTATG 838
 Db 781 TTGCTTACAGAGCGGCTTGTATCAAGACCGTGTGACGCAAAATAGGTAAGTATG 838

OY	653	GGTCAATTCGGCTCCCTTGGAATACGTCGGACACAGATGCTCAGTTGGGTCGTTT	712
Db	58746	GGCGTCTCCGTTTATATGTGGAAATATGCGCGTAACCTGAAG-----TTGAAATTC	58799
OY	713	GGTGGCTTCAATTCATGGGGCAATTCCTCTCTTACTATGTGATCATCGGATTTTG	772
Db	58800	TTTGGGATTTATATACAGAGGGCAGCCCTTGTTCGCATATTAATGSGTGCTTTC	58858
OY	773	ATGATGGATTGATCTTCAACGCGGTTTGATCAAGCCGTGTGACCAAAATAGAGT	832
Db	58860	ATTATGGCTTGGGCTTATTCACCC-----AAAAGTCCGTATATAAATAGAG	58907
OY	833	AGTTAGGTGAACAGTATTTAGATCTTTGTCAAGCGCATGTGCGATCAAGGTGTTGGCTT	892
Db	58908	ATTTTTATGAAGCAATATCTTTGAGCTTTGTGCGCGATGTGTAGTGAAGGAAATGGCTT	58967
OY	893	GAATATGAAGCAACGGGCAACGCTGTGTTGACGTGATTAATGCCATTTGACCTACGAT	952
Db	58968	GCTATATACGTCACAGGTAGCATTTGCCCTCACGCTATTAATGACATTTAGAAATATGAT	59027
OY	953	GTCGGCAACATCAGTTTCTCTAGTGACTACACGACAGATTTTGGAAAGCTGCCGTA	1012
Db	59028	GTGGCGAATATCAATTTCCGCTGATTAATCCCGTAAAGATTAATGGAACCGCGATT	59087
OY	1013	GCGAGTTGCTCGGCTATATTCGTGTATCGATTAATGCGCGGATTTTGCCAAATAGGT	1072
Db	59088	GCTGAATTTTAACTTATATTCGTGATATGACATATCCGCTGATTTCCGCCACTGGC	59147
OY	1073	ACCAAAACCTGGATGCTAATGCGAATTTAAACAGCATGGCTCAACATCTTACCGT	1132
Db	59148	ACGAAACTTGGATGCTAATGCGAATGAAATGACAGCTTGGCTTGCATATCCGATCGT	59207
OY	1133	AAAGGTGAGGATTAACATGGGAGCGGTATGAGTTCAAGGTATAGGCTTGGGCTAAGCT	1192
Db	59208	AGAGGCTTATATGATATGGGCGCGTATATGTTGTGCAAGGACAGCATAGCGTAAAGCT	59267
OY	1193	GATGTTGTCTATTTGACAGATTGAAAAAGATTGTATGATTTGAGCCGTGCGTGAAT	1252
Db	59268	AATGAGAAACATATGATGATGAGTACGTAAGTTGTTAATACCTTAACGAAAGGTATGAT	59327
OY	1253	GACCGAGTGAATTTCTTAATCTTCAATCCGGGTGATTTCAATGGGGTGTTTGGCC	1312
Db	59328	GATGAGAGAAAGATTTTAACTTTTAAACCTCGGGAAATTTGATCTTGTTGCTTCGT	59387
OY	1313	CCTTGATGTACAGCATCATCTTTTCATGTGCTGGGGATACCTTGATCTCAACAGTACT	1372
Db	59388	CCTTGATGTACAGCATCATCTTTTCCTGTGGGGATACCTTACATTTTACTAGCAT	59447
OY	1373	CAGCGTTCAATGATATGCTTGGGCTTGAAATTTCAACATGTGTGACGTTTATGTGTC	1432
Db	59448	CAGGTTCCGTATATGTTCGCTGTGATTTGAATTTCAACAAATTCAGGTGTTTACCTTC	59507
OY	1433	CTTGCGGTATGTGCACAGATACACAGGGAAGAAAGCGGCGTTGCGGATCACAGATGCTC	1492
Db	59508	TTAGCAATTATGTCTCAATCAGGCGCAAAAAGCCGCGAAGCATATCATAAATTTGTC	59567
OY	1493	AATGCGCAATTTACCAAGATCACTGGAATTTGATGTGCGGATGTGACGTTAAACGTGAG	1552
Db	59568	AATGCGCATTTTATGAAGATGAGCTGTGATTAATGCGGAGTCAATTTAAAGCGGAG	59627
OY	1553	CCATTTCCACGCGCTCAGTTCCATCATCTCAATCCAAAGATTTAAACAATCAGGATTTGGA	1612
Db	59628	CCTTTCCCATTAACCAAAATTTAGAAATTAATCCAGATTTAAAAACGCTTGAATTTGGA	59687
OY	1613	ACTTGGGTACTTTGATGATATTTTGAGCTGACCGGATATCAGTTCCAGATCTTATTCAA	1672
Db	59688	ACTTGGGTACAGATGATATTTTAAAGCTTGGCTATCATCCACAGAACCAATTTAA	59747
OY	1673	TACCGTTTCACTTAAT1691	
Db	59748	TATCTTTTCCGTTTAAT59766	

Query Match	15.9%	Score 462.2;	DB 7;	Length 852;
Best Local Similarity	71.4%	Pred. No. 4.2e-136;		
Matches 608; Conservative	0;	Mismatches 243;	Indels 0;	Gaps 0

QY	840	GGAAACGATATTGATCTTGTGACGACATCGTCGATCAAGAGTGTGGTTGAAATG	899
Db	2	TGAAGCAATATCTTGAGCTTTGTGCGCCGCAATGTTAGTAAGGGGAATGGGTTGTAATG	61
QY	900	AACGAACGGGCAACGCTGTTGTTGACTGTGATTAATGCCGATTGACCTACATGTGGGCA	959
Db	62	AACGTACAGGTAAACATATGCTCACGGCTCATTAATGACGATTTGAAATATCATGTGGGCA	121
QY	960	ACATTCAGTTCCCTCTAGTGACTACACGCAAGATTTTGGAAAGTCGCCGACGAGT	1015
Db	122	ATAATTCATTTTCCGCTGATTTCTACCCGTAAAGTTATTGGAAAGCGCGATTTGCTGAT	181
QY	1020	TGCTCGGCTATATTCGTGCTTACGATAATGCGGCGGATTTTCGCCAATTAGGTACCAAA	1079
Db	182	TTTATGTTATATTCGTGATATATGACATGCGCGTGAATTCGCGGCACTTGTGCAAA	241
QY	1080	CCGTGGATNGCTAAATGCCAATTTAAACAGATGGCTCAACATTCCTTACGCTAAAGGT	1139
Db	242	CTTGGAGATCTAAATGCGATATAAAATGACGCTTGGCTTGCAAAATCCGATGTAAGGCG	301
QY	1140	AGATGACATGGACGCGGTGATGATGTTCAAGGTAGAGCTTGGGCTAAGCTGATGTG	1199
Db	302	TTGATGATATGGGCGCGTATATGATGATGTCAGGACAGGACATGCGTAAAGCTTATGAG	361
QY	1200	GTCATATTTGACCAATTTGAAAAAGATTTGTATGATTTGAGCCGTCGGCTGATACCGAG	1255
Db	362	AAACTATGACAGCTACGTAAATTTGTTATTACTTAACGAAGAATTTATGATGATGAG	421
QY	1260	GTGAATTCCTTAACCTTACATTCGAGTCGAGTGAATTTCAATGAGGATGTTTGGCCCTTGCA	1319
Db	422	GAGAGATTTTAACTTTTTTTAACTTCGCGGAAATTTGATCTGTGTCCTTGTCTTGTGTA	481
QY	1320	TGTCACAGCAATTTTTCATTTGCTGTGGGAGTAACTGTATCTCAACAGTACTAGCGCT	1379
Db	482	TGCAATACGCAATCTTTTCTCTTGTGGCCGATCTTACATCTTACTGCTATACAGCTT	541
QY	1380	CATGTGATGTGCCCTTGGGGTGAATTTCAACATGTGTGACAGTATTATGTGCTTGGGC	1439
Db	542	CTGTGATGTTCGCTTGATTTGAATTTCAATCAATTCAGTGTTTAACTTCTTAGCAC	601
QY	1440	TGATGCGACAGATTCACAGGAAAAAGCGGCTTGGGGTATCACAAATCGTCAATGGC	1499
Db	602	TTATGGCTGCAGTACGCGGCAAAAAACCGGCAAGCAATACATPAAATTTGTGATGTGC	661
QY	1500	ACATTTACAGATCACTCGAATTTGATGCGGATGTGACGCTPAAAACTGAGGCATCTCC	1559
Db	662	ATATTTATGAAGATCAGCTTGATTAATGTGTGACGTAACTTAAACGAGCGCTTCC	721
QY	1560	CAGCGCTCAGTTCATATCATTCCTCAAAAGATTTAAACACTGACAGATTTGAAAACCTTGGG	1619
Db	722	CATTACCAAAATTTGAATTAATCAATATAAAAAAGCTTGAAGATTTGGAAAACCTTGGG	781
QY	1620	TCACTTTGATGATTTTGAAGCTCACCGGATTCAGTTCCAGATCTTATCAATACCGCT	1679
Db	782	TCAAGATGATGATTTTAAAGTCGTTGGCTATCAATCCACGAACCAATTAAATATCTT	841
QY	1680	TTTCAATCTTAA 1690	
Db	842	TTTCCGCTTAA 852	
RESULT 10			
ACA44455			
ID ACA44455 standard; DNA, 849 BP.			
XX ACA44455;			
XX AC			
XX DT 19-JUN-2003 (first entry)			
XX DE Prokaryotic essential gene #26112.			
XX AN Antisense; ds; prokaryotic essential gene; cell proliferation;			

KM drug design; gene.
XX
OS Proteus sp.
XX WO20027183-A2.
XX
XX 03-OCT-2002.
XX
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-029926/02.
XX P-PSDB; AB040585.
XX
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 32325; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 849 BP; 263 A; 155 C; 189 G; 242 T; 0 U; 0 Other;

Query Match 13.9%; Score 403.8; DB 7; Length 849;
Best Local Similarity 67.3%; Pred. No. 1.9e-117;
Matches 570; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

QY 840 TGAAGAGATTTAGATCTTTGTCAGCGCATCGTCGATCAAGGTTTGGGTGAATG 899
DB 2 TGAAGAGATCTGCGATGTGTCAACGATTATGATGACGATCAAGGTTGAATG 61
QY 900 AACGAACGGGCAAGCGTGTGTTGACTGTGATTAAATGCGGATTGACCTAAGATGTGGCA 959

DB 62 AACGACAGAACCCGTTGTTAAGGTGATTAATGCCGATCTGGAATATGATGTTGCCA 121
QY 960 ACAATCAGTTTCTCTAGAGACTACACGCAAGAGTTTGGAAAGTCCGTAGCCGAGT 1019
DB 122 ATACCAATTTCCATCATCAACACGCGTAAAGTTTTCACAAAGCCGATTCGGGAC 181
QY 1020 TGCTGGGCTATATTCGCGTTAGATATAGCGGGATTTTCGCCATTTAGTACCAAA 1079
DB 182 TATTAGGTAATGAGTTAGTTATGATATACAGACATTTTCGTATTGCTGTAT 241
QY 1080 CTTGGGATGCTAATGCCAATTTAAACCAACATGCTCAACAAATCTTACGTAAGGTG 1139
DB 242 CGTGGAACGGGAATGCTAATGAAATAGCCGATGTTAAATTAATCCGATGTAGAGGTG 301
QY 1140 AGATGACATGGGACGCGTGTATGTTTCAAGGTAGAGCTTGGGCTAAGCTGATGTG 1199
DB 302 AAGATGATATGGGCGGAGTGTATGTTTCAAGACGCCAATGCAACGCCCTGATGGCT 361
QY 1200 GTCATATTGACCAAGTTGAAAAAGATTGTGATGATTTGACCCGTGCGTATGACCGAG 1259
DB 362 CGCATTTGATTCATTTACGTAAGTGTGATATTAATTAACGATTTAGACATGTG 421
QY 1260 GTGAATTTCTTAATCTTCAATTCGAGTAAATTTCAATGGGATGTTGCCCTTGCA 1319
DB 422 GTGAGATGACATCTTTTATATATCCGAGAAACAGCGTTAGGGTGTATGCTTCATGCA 481
QY 1320 TGTACAGCATCATTTTCAATGTCGGGGATATCCTTGATATCAACAGTACAGGCTT 1379
DB 482 TGCTACATCAATCTTTTCAATGTCGGTACAGCTTATTTAATCTTCAATCAAGTAT 541
QY 1380 CATGTGATGTCCTTGGGTTGAATTTCAACATGTCAGGTTTATGTTCTTCCGCG 1439
DB 542 GCTGTGATGTCCTTGGGTTGAATTTCAACATGTCAGGTTTATGTTCTTCCGCG 601
QY 1440 TGAATGCAATGTCAGGAAAGGAAAGCGCGGCTTGGGTTATCAAGATGCTCAAGCGC 1499
DB 602 TGGTGGTCAATTCACGCGGCAATTAACAGGTAAGATTCATGATGTAATGCTC 661
QY 1500 ACATTTACCAAGATCAATCGAATGATGCGCATGTCAGCTTAAACGAGACCATTC 1559
DB 662 ACATTTAGGAACCATTTACCTGTATGCTGATGTTCACTTAAAGAGACCATTC 721
QY 1560 CAGCGCTCAGTTTCATATCAATCAAGATTTAAACATGTCAGGATTTGAAACTTGG 1619
DB 722 CATTAACCTAAGCTAATATTAATCAAAACATTAAGCATTAAGATGATGAATGAG 781
QY 1620 TCATTTGATGATTTTGAAGTCACCGGATATGATGTCAGATCCATTCATTAACCGGT 1679
DB 782 TCACGACATGATGATTTAGTTGAGGTTATGATGATGATGATGATGATGATGATG 841
QY 1680 TTTTCAGT 1686
DB 842 TCACGGT 848

RESULT 11
AA237118
ID AA237118 standard; DNA; 807 BP.
XX AA237118;
XX
XX 28-JAN-2000 (first entry)
XX
XX Nucleotide sequence of H. influenzae HI0904 gene.
XX
XX
XX Genome; mutagenesis; transposon; isolate; locate; essential gene; ss;
XX detect; growth; anti-microbial therapy; genomic footprinting.
XX Haemophilus influenzae.
XX
XX W0950402-A1.
XX


```
WP ACF67367_54 5400001 5510000
WP ACF67367_55 5500001 5610000
WP ACF67367_56 5600001 5648894

Query Match
Best Local Similarity 61.1%; Pred. No. 4.8e-76;
Matches 544; Conservative 0; Mismatches 278; Indels 69; Gaps 3;

14 ATGCTCAGAGGTTATGCAAGTTCCCAATATTGACCCGATGTTGTTTCATGCGCCCT 73
Db ATGAGTACAGCACTACCTGCGCATTTCTAATATTGACCCGATGTTTCATGAGTCCA 94398
74 CTAGCGGTGGCGGTGATGCGCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 133
Db ATAGCCCTTCATTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 94397
134 GCCAATGCGCGAGCGGATGCGCGGCGAGTGTGGACGCGTGAAGTCTCTGACTTG 193
Db GCGACACGACGAGGCGCAACCCCAATAGTGTGGCAAAAACGAAGTAGAAGAACTTG 94278
194 TTAATGCGCGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 253
Db CTTATGACAGGTTTCTGCGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTAT 94218
254 AATTTGATCTGTTCTCTGCTGACCCCTTTATTTATTTCAAAAGTGAAGTGGCGCATG 313
Db AATTTCCCTGCTTTCTGCAACAACCCCTTTATTTGTTTAAAGTTGGGATGTTGGATG 94217
314 TCTTTCAAGCGCGCTTATTTGGGTGATGACCGCATTTCTGTTGTTGCGGTAAAC 373
Db TCTTTCAAGCGGTGTTAGTGGGGTTATCTGCGCAATGTTGTTGTTGTTGTTGTTGTTG 94157
374 CAACGACCTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 433
Db AAACGCACTTTCTTAAAGTGTGACTTCACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 94097
434 ATGGAAGTATCGTAATCTTATGAATAGTAACTTTGGGACGAATAC--GGATGTG 490
Db ATGGAAGTATCGTAATCTTATGAATAGTAACTTTGGGACGAATAC--GGATGTG 94037
491 CTTGGGGCTTTGTTATCCCTA----- 512
Db CTTGGGGCTTTGTTATCCCTA----- 93977
513 -----ATGGTGGCCCATCGCCCGCCATCCCTTACAGGCTT 547
Db CCCTCACTGTATACAGTACTGTAGCAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 93917
548 TATGAATTTGGCTTAAAGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 607
Db TATGAATTTGGCTTAAAGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 93857
608 CCTGCGCCCTAGGCGAGCGTATCCGAGCTTTTATGAGTGAATGCGTATTCGCGCTC 667
Db CCTGCGCCCTAGGCGAGCGTATCCGAGCTTTTATGAGTGAATGCGTATTCGCGCTC 93797
668 CTTGTGAATAGTCCGTGATGACCAAGTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 727
Db ATAGTGAATAGTCCGTGATGACCAAGTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 93737
728 ATGGAAGTATCGTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 787
Db ATGGAAGTATCGTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 93680
788 TACAGCGCGGTTTGTATCAAGCCGTGTAGACGAATATAGGTTAG 838
Db TATTAACACCAAGTATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 93620

RESULT 13
ACF65386_0
WP Sequence split into 7 fragments LOCUS ACF65386 Accession ACF65386
WP Fragment Name Begin End
```

```
WP ACF65386_0 1 110000
WP ACF65386_1 100001 210000
WP ACF65386_2 200001 310000
WP ACF65386_3 300001 410000
WP ACF65386_4 400001 510000
WP ACF65386_5 500001 610000
WP ACF65386_6 600001 700779
ID ACF65386 Standard; DNA; 700779 BP.
ACF65386;
20-NOV-2003 (first entry)
Photorhabdus luminescens nucleotide sequence #39.
Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
detection; food; gene expression; plant; animal; microorganism; toxin;
antibiotic; biopesticide; virulence factor; disease model; plague;
whooping cough; gene; ds.
Photorhabdus luminescens.
W0200294867-A2.
28-NOV-2002.
07-FEB-2001; 2001FR-00001659.
07-FEB-2002; 2002WO-IB003040.
(INSP ) INST PASTEUR.
(CNRS ) CNRS CENT NAT RECH SCI.
Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A,
Buchrieser C,
WPI; 2003-148459/14.
Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
useful e.g. as therapeutic antimicrobials and agricultural pesticides.
Claim 1; SEQ ID NO 39; 1205bp; French.
The invention relates to the isolation of genes and their encoded
proteins from Photorhabdus luminescens. The isolated sequences are
sources of probes and primers for detecting the genome of P. luminescens
and related species; to study polymorphisms; for gene analysis and for
detection/amplification of the genes. Antibodies (Ab) raised against the
polypeptides encoded by the genes are used for detection/identification
of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
carry a gene-containing vector are used to select compounds that
modulate, regulate, induce or inhibit expression of the genes in plants,
animals or microorganisms other than P. luminescens and are able to alter
response or sensitivity to toxins and antibiotics produced by P.
luminescens. Cells transformed to express the genes are useful for
recombinant production of the proteins, particularly toxins and
antibacterial materials useful as insecticides, bactericides and fungicides. The
genes, proteins, vectors containing the genes and Ab are also useful
therapeutically (to treat microbial infection by bacteria or fungi that
are sensitive to P. luminescens-encoded toxins or antibiotics) and as
biopesticides. Other uses of the genes and the proteins are as virulence
factors and for identifying targets of human diseases for which P.
luminescens is a model (particularly plague and whooping cough). This
sequence represents one of the isolated P. luminescens genes.
Sequence 700779 BP; 190440 A; 140977 C; 164444 G; 204917 T; 0 U; 1 Other;
Query Match
Best Local Similarity 61.1%; Pred. No. 4.8e-76;
Matches 544; Conservative 0; Mismatches 278; Indels 69; Gaps 3;
14 ATGCTCAGAGGTTATGCAAGTTCCCAATATTGACCCGATGTTGTTTCATGCGCCCT 73
```

DB 104168 ATGAGTAACAGCTACTGCGATTCTCTAATATTGACCCGCTCATTTTTCATAGTCCA 104227
 QY 74 CTAGCGGTGCGCTGTATGCGCTTATGATATTTGGTGGTTCCCTTTTCTATGTTG 133
 DB 104228 ATAGCCCTCATTTGTATGTTTCACTTGTGGTTGTATTCGCCATGTTGGCTA 104287
 QY 134 GCCAATGCGGAGCGGATCGCGGCGAGTGTGACCGCTGAGCAATCTCTACTTG 193
 DB 104288 GCCACACGACGGGACGAGAACCAATAGTGTGACAAAAACGAGTATAGAACTTG 104347
 QY 194 TTATGCGCGGCTTTTAAAGTGTAGATCGGTGCGGAGTTGGTTATGATCTTTCAC 253
 DB 104348 CTTTATGACAGTTTCTGCGCATTTGTGTGGGCGCTGCTGCGCTACCTCTGTTTAT 104407
 QY 254 AATTGTATGCTTCTCTGCTGACCCCTTTTATTTATCAAGTGTGACCTGGCGCATG 313
 DB 104408 AATTCCCTGCTTCTCTGACCAACCCCTTTATTTTATTAAGTTGGATGTTGGATG 104467
 QY 314 TCCCTCAGCGCGCTTATTTGTTGTTGATACCGGCATGTTCTGTATGCGGTAATAAC 373
 DB 104468 TCTTTCAGCGGTGTATGTTGGGGGTATCTGCGCATGTGGTGGTTGGCGGAGAACG 104527
 QY 374 CAACGCACTTCTTTGTTGTTGGCGCATTTTGTGCCCCCTTATGTCATTCGTTGGG 433
 DB 104528 AAACGCAATTTCTTACAGTTGCTGACTTCACTTCCCGTATGCTCTTGGTTTGGC 104587
 QY 434 ATGGAAGTATCGGTACTTTATGATATGATGACCTTTGGGACGAGTAAC--GGATGTG 490
 DB 104588 ATGGGAGAAATCGGTAAATTTATTAACGAGTTGTGGGACGTGTCACTATGATACG 104647
 QY 491 CTTGGGCTTTTATTCCTTA----- 512
 DB 104648 CCGTGGCGCATTTATTTCCCAAGCTCCGTTGTAAGATATTGCAATGCTGCAACCAT 104707
 QY 513 -----ATGNGGCCACTGCGCGCATCTTCAACAGCTT 547
 DB 104708 CCGTCACTGTATCACTTGTGACATATGTTGTTCTGCCAGTCACTTCAACAGCTT 104767
 QY 548 TATGAATGCGCTTAAAGCGTGTGTTCTTCTTATTTCTTATTTGTTATTTGGTAA 607
 DB 104768 TATGAATGCGCTTAAAGCGATCGTTCTTCTTATTTGATTTATTAATATTCGGTAA 104827
 QY 608 CCGTCCCGCTAGGACGCGATCCGAGCTGTTTATGCTGATACGCTACCTTCCGCTTC 667
 DB 104828 CCGTCCCGCAATGGGCAAGTGTTCGGTTATTTCTTATTTGCTACGATATTTCCGAGTG 104887
 QY 668 CTTGTGAATACGTCGCTGAGCGAGATGCTCACTTGGGTCTGTTGGGCTTCATTTCA 727
 DB 104888 ATAGTGAATTTCTCCGCCAACAGATGACAACTGGCTTATTCGACGGGATCAGT--- 104944
 QY 728 ATGGGGCAATCTCTCTTACTATGTTGATCATGCTATTTTGAATGTTGGTGTCT 787
 DB 104945 ATGGGGCAATTTCTTCTATACCAATGATCTGCGCGGAAATTAATGATGATGGGCA 105004
 QY 788 TACAAGCGCGTTTGTATCAAGACCGTGTAGACGAAATAGGTAATG 838
 DB 105005 TATTAACACCAAGTATTAAGTATCAAGAGTAAATGAAACGATATCTGG 105055

RESULT 14
 ACF65386_1
 Continuation (2 of 7) of ACF65386 from base 100001 (Photobhabdus luminescens nucleotide
 WP Sequence split into 7 fragments LOCUS ACF65386 Accession ACF65386

WP	Fragment Name	Begin	End
WP	ACF65386_0	1	110000
WP	ACF65386_1	100001	210000
WP	ACF65386_2	200001	310000
WP	ACF65386_3	300001	410000
WP	ACF65386_4	400001	510000
WP	ACF65386_5	500001	610000
WP	ACF65386_6	600001	700779

Query Match 9.6%; Score 278.2; DB 7; Length 110000;

Best Local Similarity 61.1%; Pred. No. 4.8e-76;
 Matches 544; Conservative 0; Mismatches 278; Indels 69; Gaps 3;

QY 14 ATGCTCAGGCTATATCGACAGTTTCCCAATATGACCCCGTATGTTTGCATGCGCCCT 73
 DB 4168 ATGATGAACAGCTACCGGATTTCTTAATATGACCCCGTATTTTTCATAGTCCA 4227
 QY 74 CTAGCGGTGCGCTGTATGCGCTTATGATATTTGGTGGTTCCCTTTTCTATGTTG 133
 DB 4228 ATAGCCCTCATTTGTATGTTTCACTTGTGGTTGTATTCGCCATGTTGGCTA 4287
 QY 134 GCCAATGCGGAGCGGATCGCGGCGAGTGTGACCGCTGAGCAATCTCTACTTG 193
 DB 4288 GCCACACGACGGGACGAGAACCAATAGTGTGACAAAAACGAGTATAGAACTTG 4347
 QY 194 TTATGCGCGGCTTTTAAAGTGTAGATCGGTGCGGAGTTGGTTATGATCTTTCAC 253
 DB 4348 CTTTATGACAGTTTCTGCGCATTTGTGTGGGCGCTGCTGCGCTACCTCTGTTTAT 4407
 QY 254 AATTGTATGCTTCTCTGCTGACCCCTTTTATTTATCAAGTGTGACCTGGCGCATG 313
 DB 4408 AATTCCCTGCTTCTCTGACCAACCCCTTTATTTTATTAAGTTGGATGTTGGATG 4467
 QY 314 TCCCTCAGCGCGCTTATTTGTTGTTGATACCGGCATGTTCTGTATGCGGTAATAAC 373
 DB 4468 TCTTTCAGCGGTGTATGTTGGGGGTATCTGCGCATGTGGTGGTTGGCGGAGAACG 4527
 QY 374 CAACGCACTTCTTTGTTGTTGGCGCATTTTGTGCCCCCTTATGTCATTCGTTGGG 433
 DB 4528 AAACGCAATTTCTTACAGTTGCTGACTTCACTTCCCGTATGCTCTTGGTTTGGC 4587
 QY 434 ATGGAAGTATCGGTACTTTATGATATGATGACCTTTGGGACGAGTAAC--GGATGTG 490
 DB 4588 ATGGGAGAAATCGGTAAATTTATTAACGAGGATGTGGGACGTGTCACTATGATACG 4647
 QY 491 CTTGGGCTTTTATTCCTTA----- 512
 DB 4648 CCGTGGCGCATTTATTTCCCAAGCTCCGTTGTAAGATATTGCAATGCTGCAACCAT 4707
 QY 513 -----ATGNGGCCACTGCGCGCATCTTCAACAGCTT 547
 DB 4708 CCGTCACTGTATCACTTGTGACATATGTTGTTCTGCCAGTCACTTCAACAGCTT 4767
 QY 548 TATGAATGCGCTTAAAGCGTGTGTTCTTCTTATTTCTTATTTGTTATTTGGTAA 607
 DB 4768 TATGAATGCGCTTAAAGGATCGTTCTTCTTATTTGATTTATTAATATTCGGTAA 4827
 QY 608 CCGTCCCGCTAGGACGCGATCCGAGCTGTTTATGCTGATACGCTACCTTCCGCTTC 667
 DB 4828 CCGTCCCGCAATGGGCAAGTGTTCGGTTATTTCTTATTTGCTACGATATTTCCGAGTG 4887
 QY 668 CTTGTGAATACGTCGCTGAGCGAGATGCTCACTTGGGTCTGTTGGGCTTCATTTCA 727
 DB 4888 ATAGTGAATTTCTCCGCCAACAGATGACAACTGGCTTATTCGACGGGATCAGT--- 4944
 QY 728 ATGGGGCAATCTCTCTTACTATGTTGATCATGCTATTTTGAATGTTGGTGTCT 787
 DB 4945 ATGGGGCAATTTCTTCTATACCAATGATCTGCGCGGAAATTAATGATGATGGGCA 5004
 QY 788 TACAAGCGCGTTTGTATCAAGACCGTGTAGACGAAATAGGTAATG 838
 DB 5005 TATTAACACCAAGTATTAAGTATCAAGAGTAAATGAAACGATATCTGG 5055

RESULT 15
 ACF69784
 ID ACF69784 standard; DNA, 876 BP.
 AC ACF69784;
 XX 20-NOV-2003 (first entry)
 DT
 XX Photobhabdus luminescens nucleotide sequence #8251.

XX	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW	detection; food; gene expression; plant; animal; microorganism; toxin;
KW	antibiotic; diophticidae; virulence factor; disease model; plague;
KW	whooping cough; gene; ds.
XX	
OS	Photorhabdus luminescens.
XX	
PN	WO200294867-A2.
PD	
XX	28-NOV-2002.
XX	
PF	07-FEB-2002; 2002MO-IB003040.
PR	07-FEB-2001; 2001FR-00001659.
XX	
XX	(INSP) INST PASTEUR.
PA	(CNRS) CNRS CENT NAT RECH SCT.
PA	
XX	
PI	Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Darchin A,
PI	Buchrieser C,
XX	
XX	WPI; 2003-148459/14.
DR	
XX	
FT	Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT	useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX	
PS	Claim 2; SEQ ID NO 8251; 1205bp; French.
XX	

Sequence 876 BP; 200 A; 177 C; 221 G; 278 T; 0 U; 0 Other;

Query Match	9.5%	Score 277.8	DB 7	Length 876
Best Local Similarity	62.2%	Pred. No. 3.2e-77		
Matches 528	Conservative	0	Mismatches 252	Indels 69
				Gaps 37

QY 1 ATGCGCAGGGATATGTCAGTTCCCAATATTGACCCCGATATGTTTCATGCGCCCT 73
D1 1 ATAGATGACAGCTACTGGCAATTCCTAATATTGACCCGGCATATTTTCAATAGGCCA 60
QY 74 CTACGGATGGCGTGTATGATGATTTGTGTGGTTTCCTTTTGCATATGGATTG 133
D1 61 ATAGCCCTCATGTGATGATGTTTCAATGTACTGGTTGGTTTGTATTTGGCATGTGCTA 120
QY 134 GCCAATCGCCGACGGGATCGCGGGGCACTGTTGAAAGCGTGAACAAGTCTGACTTG 193
D1 121 GGCACACGCAAGGACACGAAACCAATAGTGGTTTGACAAATAACGAATGAGAACCTTG 180
QY 194 TTATTGCGCGGCTTTTAGTGTAGTGAATCGGTGGCGAGTTGTTATGATCTTCTAC 253
D1 181 CTTTATGCAAGTTTTCCTGCGATTTTGTGTGGCCCTCTGGGCTGACGTCCTGTTTAT 240

```

QY 254 AATTGATGATGCTTCCTGCTGACACCTCTTTATTTATTCAAAGTGTGACCTGGCGCATG 31.3
Db 241 AATTTCCTGCTCTTCTGCAACAACCCCTTTATTGTTTAAAGTTGGAGTGGTGGATG 30.0
QY 314 TCCTTCACAGGCGCGCTTATTTGGTGTGTGATCAACCCGCAATGTTCTGTATGCGCGTAAAC 37.3
Db 301 TCCTTCACAGGCGGTATTAGTGGGCGGTATCTGCGCAAAATGTGTGTGGCGCGAGAACG 36.6
QY 374 CAAGCAACCTTCTTTGTGTGTGCGCCGATTTGTGTGCCCTTTAGTGCATTCGGTTTGGG 43.3
Db 361 AAAGCGCATTTCTTACAGATGCTGTACCTCAATGTGCCCGGTATGTCCTTTTGTGTTAGGC 42.0
QY 434 ATGGGAGCGTATCGGTACTTATGTAAATGTACAACTTTGGGACGACAGTAAAC---GGATGTG 42.0
Db 421 ATGGGAGGAATCGGTAAATTTATTTAAACGCGAGTTGTGGGACGTGTCAATGGATGACG 48.0
QY 491 CTTGGGCTTTTGTATTCCTTA-----51.2
Db 481 CCGTGGCGCATGTATTTCACAGACTCCGATGTGAATATGTGATTTGGCTGCACCGAT 54.0
QY 513 -----ATGTGGCCCACTGGCGCGCCACTTCCTTAACAGTT 54.7
Db 541 CCTCACTGTATCACTACTTATGACGAAATAGTGTCTGCGACGTCAACCTTACAGCTT 60.0
QY 548 TATGAATTTGGCTTACAGAGCGTGTCTGTCTTTATCTTAAATGTTTATTTGGTAA 60.7
Db 601 TATGAATATGGCTGTGAAGAGATCGTCTGTCTCAATTTTGAATTAATATATCCGTA 66.0
QY 608 CCTGTGCGCTATAGCAGCGTATCCGGACTGTTTTAGTGGATACGTTACATCCGCTTC 66.7
Db 661 CCTGCGCCAAATGGGCGAGTGTTCGCGTTATCTTAAATGGCAGTAAATTTTCCAGTG 72.0
QY 668 CTGTGTGAATACGTCCGTGAGCGCAATGCTCAAGTGGCTGTGTTGTGTGCTTCATTCA 72.7
Db 721 ATATGTGAATCTTTCGCGCAACAGATTCACAACCTGAGCTTATTCAGACGGAGACAGT--- 77.7
QY 728 ATGGGCAATCTCTCTCTTACTTATGTGTATCAATCGTATTTTGAATGATGTGTTGCTC 78.7
Db 778 ATGGGCAATCTTCTTATACAAATGATCTGTGGGGAATATTTAATGATGATATGGCA 83.7
QY 788 TACAAGCG 79.6
Db 838 TATTAACAC 84.6

```

Search completed: July 31, 2004, 13:34:23
Job time : 770 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 12:39:50 ; Search time 150 Seconds

(without alignments)
10762.348 Million cell updates/sec

Title: US-09-700-712A-1

Perfect score: 2909

Sequence: 1 ggaaggttctgtatgcctc.....caggaatcgcgctgcag 2909

Scoring table: IDENTITY NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	744.2	25.6	1830121	4	US-09-557-884-1
2	744.2	25.6	1830121	4	US-09-643-990A-1
3	410.6	14.1	873	4	US-09-543-681A-2282
4	292.2	10.0	807	3	US-09-277-565-17
5	282.4	9.7	954	4	US-09-252-991A-15221
6	282.4	9.7	954	4	US-09-252-991A-15221
7	282.4	9.7	954	4	US-09-252-991A-15221
8	265.8	9.1	831	4	US-09-543-681A-2365
9	265.8	8.8	879	4	US-09-489-039A-2965
10	208.6	7.2	834	4	US-09-328-352-3810
11	152.2	5.2	640681	4	US-09-790-988-1
12	143.8	4.9	891	4	US-09-540-236-663
13	142.2	4.9	96109	4	US-09-596-002-35
14	142.2	4.0	435	4	US-09-252-991A-15244
15	75.4	2.6	2288	4	US-08-935-433-1
16	75.4	2.6	2288	4	US-09-553-132-1
17	75.4	2.6	4137	4	US-09-499-964-2
18	71.6	2.5	2058	4	US-09-252-991A-15243
19	64.8	2.2	942	4	US-09-347-878-29
20	64.8	2.2	1536	3	US-09-089-195-1
21	64.8	2.2	1536	4	US-09-367-007C-38
22	64.8	2.2	1536	4	US-09-362-663-6
23	64.8	2.2	1536	4	US-09-563-333-6
24	61.2	2.1	978	4	US-09-134-000C-941
25	57	2.0	978	4	US-09-107-532A-2491
26	54.2	1.9	7685	4	US-09-221-017B-1092
27	49	1.7	969	4	US-09-134-001C-357

28	48.2	1.7	870	4	US-09-252-991A-15224	Sequence 15224, A
29	47.4	1.6	801	4	US-09-489-039A-2884	Sequence 2884, Ap
30	46.6	1.6	867	4	US-09-107-532A-2355	Sequence 2355, Ap
31	46.6	1.6	1014	3	US-09-230-637-1	Sequence 1, Appl
32	46.6	1.6	35100	2	US-08-770-379-17	Sequence 17, Appl
33	46.6	1.6	35100	3	US-08-757-669A-17	Sequence 17, Appl
34	46.6	1.6	35100	4	US-09-230-371A-17	Sequence 17, Appl
35	45.8	1.6	5924	4	US-08-956-171E-130	Sequence 130, App
36	45.4	1.6	2335	4	US-08-961-527-275	Sequence 275, App
37	44.8	1.5	852	4	US-09-134-001C-2654	Sequence 2654, App
38	44.8	1.5	870	4	US-09-134-000C-1317	Sequence 1317, Ap
39	42.6	1.5	621	4	US-08-935-433-3	Sequence 3, Appl
40	42.6	1.5	621	4	US-09-583-132-3	Sequence 3, Appl
41	42.4	1.5	1839	4	US-09-203-895-2	Sequence 2, Appl
42	40.6	1.4	7218	1	US-08-232-463-14	Sequence 95, Appl
43	40	1.4	1568	4	US-08-936-165A-95	Sequence 293, App
44	39.8	1.4	4214	4	US-09-221-017B-293	Sequence 14, Appl
45	38.6	1.3	7218	1	US-08-232-463-14	

ALIGNMENTS

RESULT 1
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 650681
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESSES:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557, 884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match
Best local similarity 66.7%; Pred. No. 1.7e-245;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;
2 AGAAGTTTGTATGCTCAGGTTATCTGACGATTTCACATATTGACCCCGTATTTT 61

```

Db      958086 ATAGAGAAATTTAGTAATTAATTTATTTCTCTCCCACTTTGATCCGATATTTT 958145
Qy      62 TCGATCGGC-----CCTTAGCGGTCGCTGCTGATAGCTTGATGATTTGGTGGCTTC 115
Db      958146 ACGCTTGCGATAGTATATATCGGCTTACGTTGATGCGCTTGATGATACCTTTTAAAGCTTT 958205
Qy      116 CTTTTCGCTANGTGTTGGCCAAATCGCCGAGCGGATCGCGGCGAGTGGTGGACGCGT 175
Db      958206 GTTTTTCACAGTTGGCTTGCGGTTCCCGCTGATACGCCAAATGCGGTGGACAGTA 958265
Qy      176 GAGCAAGTCTCTGACTGTTTATTGCGCGCTTTTATGAGTGAATGATCGGTGGCGAGTT 235
Db      958266 GATCAAGTGAAGTACTTACTTTTCAACGCTTTTATGGGCGTGTATTATGGCGGACGTGAT 958325
Qy      236 GGTATGATGATCTTCTCAATTTTGAATGATGCTGCTGCTGACCGCTTTTATTTATTCAMA 295
Db      958326 GCGCATGATTTTCTTATATCTCGATCATTTCTTACAGAACCACTTTTATTTATTCGCG 958385
Qy      296 GTGTGACTGCGGCGATGCTTCCACAGCGCGCTTATGGGTGATTCACCGCATGTTTC 355
Db      958386 GTTTGGAGAGTGAATGATGCTTCCACGCTGCTTATGTTGTTATGTTGCTATGATTT 958445
Qy      356 TGTATGCGCGTAAAAACAAGCACTTCTTTGCTGCTGCGATTTTGTGCGCTTTG 958505
Db      958446 TGGACATCTTATCTCAAAAACGTAATTTTGGCAACGCTGATTTTGTGCGCTTTG 958565
Qy      416 GTGCGATTCGCTTTGGGATGGGACGATCGTAACTTTATGATAGTGAATCTTTGGGA 475
Db      958506 ATTCCGTTTGTTTATGTTTATGAGAGAAATGTTATTTATATCTGATGACATGAGGGA 958565
Qy      476 CGAGTAACGATGCTGCTTGGGCTTTTGTATTCCTTATAGTGGCC---CACTGCGCGCG 532
Db      958566 CGCGAAACGATGCTTGGGCAATGATTTCCCGATGATCTCTTTTACGCGCTGAT 958625
Qy      533 CATCTTCACAGCTTTATGATTTGCTTGAAGCGTGGTCTGCTTCTTATCTTAAAT 592
Db      958626 CATCATCAACAACCTTTATGAAGCTTTTGAAGCGCTGCTGTTTATGATCTCGAAT 958685
Qy      593 TGGTTTATGTTAAACCTGCTCGCTAGGACGATTCGCGACTGTTTATAGCTGATAC 652
Db      958686 ATTTTATTAATAAACACGCTCCATGCTTCTGTTGACGTTTATTTCTTAAATGTTAT 958745
Qy      653 GGTACATTCGCTCTCTTGGGATATAGTCGCGTAGGCAATGCTGATGGTCTGTT 712
Db      958746 GGGCTCTTCGCTTTTATTTGGAATATGCTGTAACCTGAG-----TTGAAATTTTC 958799
Qy      713 GGTGCTTCATTTCAATGGGCAATCTCTCTTACCTATGATGATCGGATTTTG 772
Db      958800 TTGGGATTTATTAACAGGCGCAAGCCCTTGTGCGGATGATTAATGGTGTCTTTC 958859
Qy      773 ATGATGTTTGTCTTTCACAGCGGCTTGTGATCAAGCCGTGACGCAAAATAGGGGT 832
Db      958860 ATTATGCTTGGGCTTATTCACGC-----AAAGTGGGCTCATTAATAATAGAG 958907
Qy      833 AGTAGAGTAAGACGATTTAGATCTTTGTACAGGCAATCTCATCAAGGTGTTGGGT 892
Db      958908 ATTTTATGAAGCAATATCTTGACCTTTGTGCGGCAATGTTGTAAGGAGGATGGGT 958967
Qy      893 GAAATGACGACAGGCGACGCTGTTTGACTGTGATTTATGCCATTTGACTTACAT 952
Db      958968 GCTATGACGCTACAGGTAAGCATGCTCATCGCTCATTAATGACGATTTAGAAATGAT 959027
Qy      953 GTGGGCAACATGATTTCTCTATGATCTACGCAAGATTTTGGAAAGTGGCGTA 1012
Db      959028 GTGGGCAATATCATTTTCGCTGATTAATCTCCGTAAGATTTTGGAAAGGCGGAT 959087
Qy      1013 GCCGAGTTGCTGGCTATTTCTGTTTACGATATGCGGCGGATTTTCCCAATTAAGT 1072
Db      959088 GCTGAATTTTATGATTTATCTGATGATGACATGCGCTGATTTTCCGCGACTTGGC 959147
Qy      1073 ACCAAACCTGGGATGCTAATGCAATTTAAACAGATGGCTCAACATCTTACCGT 1132

```

```

Db      959148 ACGAAAACCTGGGATGCTAATGCGAATGAAAATGACAGCTTGGCTTGCAAAATCCGATCGT 959207
Qy      1133 AAGGTGAGATGATGACATGGGACGCGTGTATGTTTCAAGGATAGAGCTTGGGTAAGCT 1192
Db      959208 AGAGCGTTGATGATATAGGGCGGTATATGCTGTGCAAGGACAGATGAGCGTAAGCT 959267
Qy      1193 GATGTGTCTATATTGACAGTTGAAAAAGATTGTTGATTAATTTAGCCGTGGCTGAT 1252
Db      959268 AATGAGAAATCATGATACAGCTACGTAAATGTTTAAATTAACCTTAAGAAAAGGTATTGAT 959327
Qy      1253 GACGAGGTGAATTTCTTAATCTTACATCCGGGTGAATTTTACATGAGGCTGTTGCGC 1312
Db      959328 GATAGAGAGATTTTAACTTTTAACTTTTAACTTGGGAATTTGATCTTGCTTGTCTGCT 959387
Qy      1313 CTTGCAATGATACAGCATCATTTTCAATGCTGGGAGTAACTTTGATCTCAACAGTACT 1372
Db      959388 CTTGATGATGACGATACATTTTCTCTTGTGGGCAATCTTATCACTTACTAGCTAT 959447
Qy      1373 CAGGCTCATGATGATGAGCCCTTGGGGTGAATTTTCAACATGGGAGGTTTATGTTG 1432
Db      959448 CAGGCTCTGTTGATGTTCCGCTTGATGATTAATTAATCAATTCAGGTTTACTTC 959507
Qy      1433 CTGCGCTGATGACAGATACAGGAAAAAGCCGCGCTTGGGATTCACAAAGATGTC 1492
Db      959508 TTAGACTTATGCTCAGATACAGGCAAAAAAGCCGCAAGGCAATCATTAATAATGTG 959567
Qy      1493 AATGCGCATTTTACCAAGATCACTGATGATGCGCATGTCAGCTTAAACGTGAG 1552
Db      959568 AATGCGCATTTTATGAAGATCACTGATGATGATGCTGATGATGATGATGATGATGAT 959627
Qy      1553 CCAATCCAGCGCTCAGTTCCATATCATCAATCAAAATTAACATCTGACAGATTTGAA 1612
Db      959628 CTTTCCCATTAACCAAAATTTGAATTAATGATGATTAATAAAGCTTGAAGATTGGA 959687
Qy      1613 ACTTGGTACTTGTGATGATTTTGAAGTCAACCGGATATCAATGCTCACGATCTTCA 1672
Db      959688 ACTTGGTCAAGATGATGATTTTAAAGTGTGCTATCAATCCAGAACCATTTAA 959747
Qy      1673 TACCGCTTTCAGTCTAAT 1691
Db      959748 TATCCTTTTCGCTTAAT 959766

```

```

RESULT 2
US-09-643-990A-1
; Sequence 1, Application us/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>

```

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07
ABPICATION NUMBER: 004

APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21

ENDING DATE: 1993-04-2

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PER

TELECOMMUNICATION INFORMATION:
TELEPHONE: 361 510 5700

TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439

EDFAX: 310-303-6433
ON FOR SEO ID NO: 1:

SEQUENCE CHARACTERISTICS

LENGTH: 1830121 base

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

DESCRIPTION:

U
U
C
E
F

match

Query Match	25.6%;	Score 744.2;	DB 4;	Length 1830121;
Best Local Similarity	66.7%;	Pred. No. 1.7e-245;		
Matches 1134; Conservative	0;	Mismatches 538;	Indels 27;	Gaps 4

QY	2	GAAGAGTTGGTATAGCCACAGGGTATACGACATTTCCCAATATGACCCCGATATGTTT	61
Db	958086	ATMAGAAATATTGAATTCAAATTATTTACTCTTCCCACTTGATTCAGATATTTT	958144
QY	62	TCGATCCG-----CCTTAGCGGTGCGCTGGTATGAGCTGATGATTTATGGTGGGTTTC	115
Db	958146	ACGCTTGGCGATAGTATATACGGCTTACGTTGATAGCGCTTATGATACCTTTTAGTTT	958205
QY	116	CTTTTGTGATGATGAGTGGCCAAATCGCCAGCGGATCGCGCGGACATGTTGACGCGT	175
Db	958206	GTTTGTGACGTGGCTGGCTGGGTGCGGTCCCGGTCAATCGCCAAATAGCGGTTGGACAGTA	958266
QY	176	GACAGAGCTCTGACTGTGTTATTCGCGCGCTTTTGGATGATGATCGGTGGCCGAGTT	235
Db	958266	GATCAAGTGAATGACTTACTTTCAACGGTTTATATGGGGGTGTTATATGGCGACGTTT	958322
QY	236	GCTTATGTGATCTTTCACAAATTTGATCTGTCTCTGTGACCCCTTTATTTATTCAA	295
Db	958326	GCGCATGATATTTTCTTAACTCGAATATTCTTTCAGAGAACCACTTTATTTATTCGC	958386
QY	296	GTCGGAACCTGGCGGCATGCTCTTCACACGCGGCTTATGGGTGTATACCGCATATGTC	355
Db	958386	GTTTGGAGAGGTGAATGTCGTCCACGGGTGTTATATGGATATGTTGTCATGATTT	958446
QY	356	TGATATCGCGGTAAAAACACACGACACTCTTTGTGTGCGCATTTTGTGCCCTTTA	415
Db	958446	TGCACATCTTATTCGAAAAAGCTAATTTTGGCAAAAGCGCTGATTTGTGGCCGCTTG	958506
QY	416	GTGCCATTCGTTTGGGGATGGAGCATTCGTAACCTTTATGAATAGTAACTTTGGGGA	475
Db	958506	ATTCGTTTGGTTTAGGTTTAGGCAAAATGGATATTTTCATTAATCTTGAACCTTAGGGA	958566
QY	476	CGATTAACGAGATGACCTTGCGCTTTTGATTCCTTAATGGGCGC---CACTCGCCGCG	532
Db	958566	CGCAAAACGATATGCTTGCGGAATGATTTTCCGATGATCTCTTTTACTGCTTCGT	958626
QY	533	CATCTTCACAGCTTTATGAATTCGCTTAGAAGCGGTGTCGTCTTTATTTCTTAAT	592
Db	958626	CATCCATCAACACTTTATGAGCCTTTTAGAAGCGCTGATGTTTATTCAGATCTGAAT	958686
QY	593	TGGTTATATGTAACCTGTCGCGTAGGACAGGATCCGACGTTTGTAGCTGATAC	652
Db	958686	ATTTTATTTAAAAACACGCTCAATGGCTTCGTGACAGGTTATTTCTTAATGGTAT	958746
QY	653	GGTACATTCGCTCTTGATGATACGTCCGTGACAGATGCTCAATGGGTCTGTTT	712
Db	958746	GGCGTCTTCGTTTATATGGGAATATGGCGTAACCTGAG-----TTGAATTTTC	958806

QY	711	GGTGGCTCATTTTCAATGGGGGCAAAACCTCTCCTTACCTTATGGGATCATCCGATTTTG	772
Db	958800	TTTGGATATTATCAAGAGGGCAAGCCCTTTGCTTGCCATGATTAATGGTGCTTTC	958855
QY	773	ATGATGGTTTGGTCTTTACAAGCCGGTGTGTATCAAGACCGTGTAGCAGCAAAATAGAGT	832
Db	958860	ATTATGGCTTGGGCTTATTCACCC-----AAAGTGCGGGCATTAATAATAGAG	958900
QY	833	AGTTAGTGAAATACGATTTTAGTCTTTTGTACGGCATGCTCATCAAGTGTTTGGGTT	892
Db	958908	ATTTTATGAGCAATATCTTGACCTTTGTGCGCATTTGTATGTAAGGGGAATGGGTT	958966
QY	893	GAATATGAACGACCGGGCAAGCCGTTTGTGACTGTGATTAATGCCGATTTGACCTACGAT	952
Db	958968	GCTAATGAACGATACAGGTAGACATTGCTCCACGGTCATTAATGCAATTAATAGAT	959022
QY	953	GTGGGCAACAATCAGTTTCTCTAGTACTACACGCAAGATTTTGGAAAGCTGCCGTA	1012
Db	959028	GTGGCGAAATATCAATTTCCGCTGATTAATCACCCTGAAAGTATTTGAAAGGGCGAAT	959088
QY	1013	GGCGAGTTCTCGGCTATATTCGTGTGTACGATATGCGGCGGATTTTGGCAATAGGT	1072
Db	959088	GCTGAATTTTATAGTTATATTCGTGATATGACAAATGCCGCTGATTTCCGCGCACTTGGC	959144
QY	1073	ACCAAAACCTGGAGTCTAATGCCAATTTAAACAAACATGCGTCAACAATCCTTAACGT	1132
Db	959148	ACGAAAACTTGGGATGCTATGCGAATGAAATGCAAGCTTGGCTGCAAAATCCGACATCGT	959200
QY	1133	AAAGGAGAGATGACATGAGGAGCGCGTATGAGTTCAGAGGTAGACTTGGGCTAAGCCT	1192
Db	959208	AAGAGCGCTTATGATATGAGGCGCGCTATATGAGTGCAGAGCAGACATGCGCTAAGCCT	959264
QY	1193	GATGTGTGATATTAATGACCAGTTGAAAGAAAGATTGTGATGATTTGAGCCGTGGCTTGAT	1252
Db	959268	AATGAGAAACTATGATCAGCTACGTAACGTAAATTTGTTAATTAATCTTAACGAAGTATTGAT	959322
QY	1253	GACCGAGTGAATTTCTTAACCTTCACAAATCCGGGTGAATTTTCAATGGGGTGTTCGCC	1312
Db	959328	GATAGAGGAGATATTTAACTTTTAAACCTCGGGAAATTTGATCTTGGTGTCTTCGT	959384
QY	1313	CCTTGATATACGCCCATATTTTTCATCTGCTGGGGATACCTTATCTCAACAGACT	1372
Db	959388	CCTTGATATACGATATCTTTTCTCTTGAGGGGATACCTTAATCTTACACTAT	959444
QY	1373	CAGCGTTATGATGATGTCCTTGGGGTGTGAATTTCAACATGATGAGGTTTATGTCTTC	1432
Db	959448	CAGCGTTCTGTGATGTTCCGTTGGATTTGAATTTCAATCAATTCAGGTGTTTACCTTC	959500
QY	1433	CTTGCGCTGATGAGCAACAGATCAACAGGAAAAAGCCGGGCTTGGCGATACAAAGATCGTC	1492
Db	959508	TTAGACATTAATGCTCAAGTACCGGCAAAAAAGCCGGCAAGCATATATATAAATTGGG	959564
QY	1493	AATGCGCACTTTTACCAAGATCAACTCGAATTTGATCGCGATGTGACGTCAAAACGTGAG	1552
Db	959568	AATGCGCATTTTATGAGATCAGCTTGATTAATGCGGAGCTACACTTAAACCGCAG	959624
QY	1553	CCATTTCCAGCGCCTCAGTTCATATTCATCCAAAGATTTAAACAATGACAGATTTGGAA	1612
Db	959628	CCTTTCCATTAACCAAAATTAATATCCAGATATAAACAACGTTGAAGATTTGGAA	959684
QY	1613	ACTTGGGTCATTTGGATTTTATGACGTCACCGGATATCACTTCCACAGATCTTATCAA	1672
Db	959688	ACTTGGGTCAGATGATTTTAAAGTCGTGTGCTATCAATCCACGACCAACCAANTTAA	959744
QY	1673	TACCGGTTTTCAGTCAAT 1691	
Db	959748	TATCCTTTTTCGTTCAAT 959766	

GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709,1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 2282
LENGTH: 873
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2282

Query Match 14.1%; Score 410.8; DB 4; Length 873;
Best Local Similarity 67.0%; Pred. No. 5,2e-132;
Matches 583; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

QY 820 AGCAAAATAGGTTAGTTAGTGAACAGTATTTAGATCTTTGTACGGCATCGTGATCA 879
DB 3 AAAAAAGAGAGACTACATGAGAGAGATCTGGCATTTGTCAACGATTAATCGATGA 62

QY 880 AGGTGTTGGGTTGAAAAATGAGAAAGGCAAGCGTGTGTTGACTGTATTAATGCCGA 939
DB 63 AGGTGAATGATGATGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122

QY 940 TTGACCTACGATGAGGAGCAACATCATGTTCTCTAGTACTACGCAAGATTTTG 999
DB 123 TCTGGAATGATGATGTTCCATTAACCATTTCCATCATCAACACCGTAAAGTTTAA 182

QY 1000 GAAGCTGCGGTACCGCATGTTGCTGCTATTTCTGGGTATGATTAATGCGGAGATT 1059
DB 183 CAAGCGCGCATGCGGAGATTAATGAGGATCTTAACGAGTATGATTAATGAGAGAGAT 242

QY 1060 TCGCATTAAGTACCAAACTGGAGTCTAATGCAATTTAAACCAAGCATGCTCAA 1119
DB 243 TCGGCTATGAGTGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302

QY 1120 CAATCTTACCGTAAAGTGAAGTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
DB 303 TAATCCCATCTGTAAGGATGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362

QY 1180 TTGGGCTAAGCTGATGAGTGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
DB 363 ATGGAAG 422

QY 1240 CCGTGGGAGTGAATGAG 1299
DB 423 TAACGATTAAGCATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 482

QY 1300 GGGGTTTGGGCGGCTGAGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359
DB 483 AGGGGTTTGGGCGGCTGAGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542

QY 1360 TCTCAAGATGATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1419
DB 543 TTAACTCATATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 602

QY 1420 GGTATATGTTCTTGGGCTGAGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479
DB 603 ATGCTTGTATTAATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662

QY 1480 TCACAGATGATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1539
DB 663 CCAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722

QY 1540 GCTAAAGATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1599
DB 723 ACTAAAGATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782

QY 1600 GAGAGATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1659

DB 783 AGATGATTAAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842

QY 1660 CGATCTTATTAATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1689

DB 843 CGAGGATTAATGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872

RESULT 4
US-09-277-565-17
Sequence 17, Application US/09277565
Patent No. 6207384
GENERAL INFORMATION:
APPLICANT: Mekalanos, John J.
APPLICANT: Akertley, Brian J.
APPLICANT: Rubin, Eric J.
APPLICANT: Camilli, Andrew
TITLE OF INVENTION: SYSTEMATIC IDENTIFICATION OF ESSENTIAL
FILE REFERENCE: 00742/052002
CURRENT APPLICATION NUMBER: US/09/277,565
PRIOR FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: 60/079,770
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 807
TYPE: DNA
ORGANISM: Haemophilus influenzae
US-09-277-565-17

Query Match 10.0%; Score 292.2; DB 3; Length 807;
Best Local Similarity 65.0%; Pred. No. 1e-90;
Matches 467; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

QY 81 TCGGCTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 140
DB 74 TACGTTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 133

QY 141 GCCGAGCGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 200
DB 134 GCGGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 193

QY 201 CCGGCTTGTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 260
DB 194 ACGGTTTATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 253

QY 261 ATGCTGCTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 320
DB 254 ATGATTTCTTACAG 313

QY 321 ACGGCGGTTATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 380
DB 314 ACGGCGGTTATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 373

QY 381 CCGTCTTGTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 440
DB 374 ATTTTGGCAAG 433

QY 441 GTATCGTATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 500
DB 433 GATTTGTAATTTCTTATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 493

QY 501 TTGTAATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 557
DB 494 TGAATTTCCGATGATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 553

QY 558 CCGTAAAG 617
DB 554 TTTHAG 613

QY 618 TAGGAG 677


```

RESULT 7
US-09-252-991A-15227
: Sequence 15227, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 15227
: LENGTH: 1454
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15227

```

QY 389 GGTGTGCGCATTTTGTGTGTCCTTTATGTCCATTTCGTTTGGGGATGGGACGTACCGG 448
Db 788 CAGCTATGACCTTATCGCCCCCGTGTGTCCATCGCCCTGGCGCGCGGACGATGGC 847
QY 449 AACTTATGATAGTAACTTTGGGGAGAGTAAACGATGTGCCTTGGGCTTTGTATTC 508
Db 848 AACTTATCAACTCGGAACTGTGGGGCAAGTACGACATGTGCCCTGGGCATGTCTTC 907
QY 509 CCTAATGTGGCCCACTGCGCGGCATCTTCAAGCTTATGTAATTGCTTTAGAAAGC 568
Db 908 CCGAAGCGGCGCGCTGCGCGGCATCCTTCGACAGCTGACCACTTGCTGCGCTGGAAGCG 967
QY 569 GTGTCTGTCTTTATTTCTTATTTGTATTGTGTAACCTGTGCGGTAGGACAGGTA 628
Db 968 GTGCGCTGTGTCTACTCTCTGTGCTGTACCGGCAAGCAGCGGCAGCGCTTGCTC 102
QY 629 TCCGCACTGTTTATAGCTGATACGATCATTTGCGCTTCTCTGTGTAATACGTCTGAG 688
Db 1028 TCCGGCTGTGTGTGTGTGCTAGAGGATCTTCGGCTTGTGTGCAATTCGTCGGGTG 108
QY 689 CCAAGATGCTCAGTTGGG--TCTGTTTGTGTGCTTCAATTCAATGGGGCAATCTCTCG 745
Db 1088 CCGGATGCGCAGCTCGGCTACCTCCGCTGGGGTTGGCTGACCAAGGCGCAGGTCTCG 114
QY 745 TTAACATGATATCAATCGATTTTGAATGATTTGATCTTACAGACCGG 797
Db 1148 GTACCAATGTGTGCGCGCATTTGCTCCTGATGTGTGGGCTTACCGCGCG 1139

	Query Match	9.1%	Score 265.8	DB: 4	Length 891	
	Best Local Similarity	61.6%	Pred. 1.7e-81			
	Matches 519	Conservative	0	Mismatches 252	Indels 72	Gaps 3
Qy	26	TATTCGAGTTTCCCAATATTTGACCCCGCATTTGTTTCGATCGGCGCTTAGCGGGTGGCG	85			
Db	19	TACCTTAATTTCCAGAAATCGATCCATTTATGTTTCTATCGGGCCAGATCTTTGGCAC	78			
Oy	86	TGCTATGCGTGAATGATTTGGTGGGTTCTCTTTTGTCTATGATGGTGCACATCGCCA	145			
Db	79	TGCTACGCGATGATGATTTGATTTGCAATTGTTGTTTCGATTAATGGCTAGCAACCCGACA	138			
Oy	146	GCGGATCGCGCGGCGAGTGTGTGACGGCTGACGAAGTCTCTGACTTGTATTTCCGGCG	205			
Db	139	GCGGCAAAACCTAACAGTGGTGGCGAAAAGTGAAGTCGAACATTACTTATTTGAGGG	198			
Oy	206	TTTTTAGTGTGATGATCGGTGGCGGATTTGGTATGATCTTCTACATTTTGTACTG	265			
Db	199	TTTGTGGGGGTGTTTATTGGTGGGGGTCTAGAGATATGTTTATCTTAATCTTACCGGTA	258			
Oy	266	TTCTCTGCTGACCTCTTTATTTATTCAAAGTGGACATGGCGGCAATGCTCTCCAGCGC	325			
Db	259	TTTTTAGATATTCATTTTACCTATTAAAGTCTGGGATGCGGAATGTCCTTCCAGGT	318			


```

QY      326  GGCTTATGGAGTGTATCAACCGCATGTTCTGTGTATGGCGGTAAACCAACGACCTTC 385
Db      319  GCGTTATTTGGGTATCTCGCCCATGATTTGGTTGCTCGTCGCATCAATACGTCATTT 378
QY      386  TTGTGTGTGGCCGATTTTGTGTGCCCTTTAGTGCATTGCGTTTGGAGATGGAGCGTATC 445
Db      379  TTTCAGATCGCAACACTTGTGTGCATTAATCCCATTTGGTTTAGTGAAGCGTATT 438
QY      446  GGTAACCTTAATGAATGTGTAACCTTGGGAGAGATAC---GATATGCCCTTGGGCTTTT 502
Db      439  GGTAACTTTATTAATGGAAGCTTTGGGAGAGATCACTTAATACCCCGTGGGCGATT 438
QY      503  GTAATCCC----- 510
Db      499  TTATCCCAAGGCTCCCGTCAGAGATCTACAGTTGTGCGCAAGATCCCATACACTA 558
QY      511  -----TAAATGGGCCACACGCGCGGCCCATCCCTTCAACAGTTTAAATTC 556
Db      559  TTACCTATCAATTAGCAATAACGCGATTTGCTCGCACCCATGCGCAACTTAATGAATG 618
QY      557  GCGTTAAGAGCGGTGTCTGTTCTTTATTTCTTAATGTATTATGTGTAACCTGTCGG 616
Db      619  TTCCCTAGAAGCGCTGTATTATTATTTATTCGAATATCTTTGTGCGCAAAATCGTCT 678
QY      617  CTAGGACGCGTACCGGACGCTTTTATAGTGAATACGTAACATTCGCTTCCTGTGGA 676
Db      679  ATGGGCGATATACGGGATTTCTTAATGTGTTTAGTGCTTCGATTAATGTGTGAA 738
QY      677  TAGCTCCGTAGGCAGATGCTCAGTGGTCTGTGTGATGAGCTTCATTTCAATGGGCAA 736
Db      739  TTCTTCGCCAACCTGATGCGCATAGATAGTTATTAGCGGATCA---GCATGGGCAA 795
QY      737  ATCTCTCTTACTTANGATGATCAACGCTATTTTGAATGATGTGGTCTTAACACGC 796
Db      796  ATTCTTTCATTCCAATGATTAATATTAGGCATTATCTTAATGTGTGGCGATGTCGC 855
QY      797  GGT 799
Db      856  GAT 858

RESULT 9
US-09-489-039A-2965
; Sequence 2965, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2965
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2965

```

Query Match: 14, Score 256.8, DB: 4, Length 879;
 Best Local Similarity: 59.5%, Pred: No. 2, 3e-78;
 Matches: 515; Conservative: 0; Mismatches: 282; Indels: 69; Gaps: 2

Ox	14	A T G C C T A G G C G T A T C T G C A G T T C C C A T A T G A C C C C A T G T T T G C A T G C G C C T	73
Db	4	A T G A T A G T G C T A C C G A C T G C A T T T C G G A G T T C G A C C G A T A T T T T C T A G G C C G	63
Ox	74	C T A C C G G T G C G C T G T A T G C T T A T A T T T G T G G G T T C C T T T T G C T A T G T G T T G	123
Db	64	G T C T C C C T C A C C G A T G C G T T G A T A T A C C T G G T G G C C T T T G C T T T G C A A T T G C T G	123

QY	13	GCACATGCGGAGCGGATGCGCGGGGACAGTGGTTGGAAGCGGTAGGACAGTCTTGACTTG	133
Db	124	GCCACCGGTGCGCGAATGCTCCGGGACAGCGCTGAGACGAAAATGAAAGTGGAAAACCTG	183
QY	194	TTAATGCGCGGCTTTTAAAGTGAATGATCGGTGGCGAGTGGTGTATGATCTTCTAC	253
Db	184	CTGTACCGGGCTTCTCCGGCTCTTCCGGGCGGCGGTATCGGCTACGTGCTGTATTAC	243
QY	254	AATTTGATCTGTCTCTTGCTACCGCTCTTATTTATTCAGAGTGGATCGGCGCATG	313
Db	244	AACCTGCCGATTTTCTTGCCGATCGCTGATCTGTTCGGGCTTGGAGACGGCGCATG	303
QY	314	TCTTTCACGCGCGCTTATTTGGGTGATCAACCGCATGTTCTGGATGCGCGTAAAC	373
Db	304	TCTTTCACGCGCGCTGATCGCGGTGATCTCGGTGATGATTAATCTTCCGACAGCGACC	363
QY	374	CAAGCAGCCTCTTTGTGTGGCCGATTTGTGTGCCCTTTAGTGGCAATGCGTTGAGG	433
Db	364	AAAGCACTCTTCCAGGTCTCCGATTTTANCGGCGCGCTGATCCGTTGGTCTTGGC	423
QY	434	ATGGAGATATGCGTAACCTTATGAAATGTAACCTTGGGAGCGATG-----	481
Db	424	GCGGGGCTGTGGGCAACTTATCAACGCGGAGCTGTGGGCGCGGTAGACCGGAGCTTC	483
QY	482	-----ACGGAT	481
Db	484	CATTACACATATTTTCCCGGGCTCCGTGCCGAAGACCTGGGCGTCTGCCGACGAT	543
QY	488	GTGCTTGGGCTTTTGTATTCCTTAATGATGGTGGCCCATGCGCGCGCATCTTCAACGCTT	547
Db	544	CCGGAATGGCAATCGCTGTTCGATACCTAACGCGCGCGCTGCCGCGTCAAGCTCCACACTG	603
QY	548	TATGAAATGCGCTTAAAGCGGTGTGCTTATCTTAAATTTGGTTAATGGTAAA	607
Db	604	TACGAGCTGGCTGGAAGAGCGTGTGATCTGTTCGATCTCAGACTGTTTATCCGTAAA	663
QY	608	CCTCGTCCGTAGGACGCTATCCGGACTGTTTTAGTGGATACGGTACATTCGCGTTC	667
Db	664	CCGCGCCCGACCGGCTCCGCTTTCGCGGTCTGTTCTCGATGAGCTACGCGCTGTTCGCAATC	723
QY	668	CTTGTGGAAATAGTCCGTGAGCCAGATGCTACGTT---GGGATGTTGGTGGCTCATT	727
Db	724	ATGTGTGAATTTCTTCCGCAAGCCCGACCGCCAGTTCAACCGCGGCTGGGTGCAATATC	783
QY	725	TCAATGGGCAATCTCTCTCTTAACTTATGATGATCGATCGGTATTTGATGATGTTGG	784
Db	784	AGCATGGGAGATTTCTTTCGATCCCGAATGGTGTTCGGGTATCATTAATGATGTTGG	843
QY	785	TCTTTCAGCGCGGTTTGTATCAAG	810
Db	844	GCTTACCGCATCGTCCGACGACAA	869

```

RESULT 10
US-09-328-352-3810
: Sequence 3810, Application US/09328352
: Patent No. 6562958
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: CFC98-033A
: CURRENT APPLICATION NUMBER: US/09/328,352
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 3810
: LENGTH: 834
: TYPE: DNA
: ORGANISM: Acinetobacter baumannii
US-09-328-352-3810

Query Match          7.2%, Score 208.6, DB 4, Length 834.

```


QY 149 GATCGCGGCGGAGTGTGACGCGTGAGCAAGTCTCTGACTGTATTGCGCGGCTTT 208
DB 253 AGCGCTTGAACCCGACCTGAGCCAGAGCGCCTTTCCGACCTGGCTCTTGGGCGCC 194
QY 209 TTAGGTGATGATGCGGCGGAGTGTGATGATGATGATGATGATGATGATGATGATG 268
DB 193 TCGGCGGTGATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 134
QY 269 CTGTGACCCCTCTTTATTATTGCAAGTGTGACGCGGCGGCGGCGGCGGCGGCGG 328
DB 133 ATCGCCAGCCGACGCTGATCTTGAAGGTGGAAGGCGGCGGCGGCGGCGGCGGCGG 74
QY 329 TTAGGTGATGATGCGGCGGAGTGTGATGATGATGATGATGATGATGATGATGATG 359
DB 73 CTGCTGGCGGTGATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 43

RESULT 15

US-08-935-433-1/C

; Sequence 1, Application US/08935433

; Patent No. 6319688

; GENERAL INFORMATION:

; APPLICANT: FILLD, JOHN

; TITLE OF INVENTION: A HUMAN SODIUM DEPENDENT PHOSPHATE

; TITLE OF INVENTION: TRANSPORTER (IPT-1)

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P. O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/935,433

; FILING DATE: 23-SEP-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/044,974

; FILING DATE: 28-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-70006

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2288 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; US-08-935-433-1

QY 2596 ATTGTCACAGTAGTACCAATATTGGACCCCATCAACCATAGGAATCGCGTTTCAACCG- 2654
DB 651 GTTGATGATGACGTTTCAATGTTGCCCCCATGATTAATGGGATGGGACCGCCGAAAGT 592
QY 2555 --GTAACCCACCGGCAACGAGACCAATATATAGAGTCAACCGCTTGAAGATTGAAT 2712
DB 591 GAGCAATGAAAGAGACACCATGCTGACACGATGACGTTAGAGTGTGAGGCTGTGAC 532
QY 2713 CAGTCCGTTGCACTTAAACCAATCATCAATCTGCAATTGGGTGGGAAGCA 2765
DB 531 CAAGACGATCACGACACCCGATCACAGCCCAACAAAGGTTGACATTA 479

Search completed: July 31, 2004, 17:02:39
Job time : 160 secs

Query Match 2.6%; Score 75.4; DB 4; Length 2288;
Best Local Similarity 56.0%; Pred.No.7.8e-15;
Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;
QY 2476 CATTTCTAGAGGTAGAGATCAGCAGCGAGAGATTGAAAAATCGTAGTGTGGC 2535
DB 771 CACCTTCAAGGAG 712
QY 2536 ACTGGGAAAG 2595
DB 711 TCTGTAAAG 652

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:39:00 ; Search time 7452 Seconds

(without alignments)
16919.604 Million cell updates/sec

Title: US-09-700-712A-1

Perfect score: 2909
Sequence: 1 gagaagattgtatgctc.....cagcaatcgagctgag 2909

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank :
1: gb_ba :
2: gb_hlg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vl :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :
29: em_vl :
30: em_hlg_hum :
31: em_hlg_inv :
32: em_hlg_other :
33: em_hlg_mus :
34: em_hlg_pln :
35: em_hlg_rod :
36: em_hlg_mam :
37: em_hlg_vit :
38: em_sy :
39: em_hlg_hum :
40: em_hlg_mus :
41: em_hlg_other :
Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2909	100.0	2909	1	VCH6514
2	2909	100.0	2909	6	AX009481
3	2870.6	98.7	9931	1	AE004153
4	1390.8	47.8	1410	1	VCH71A6
5	1222	42.0	1222	6	AX009483
6	1216.8	41.8	1236	1	AY143429
7	1187	40.8	1591	1	VCA10968
8	1052.4	36.2	298900	1	AP005074
9	1035	35.6	301442	1	AE016798
10	1028.6	35.4	249150	1	AP005332
11	841.4	28.9	852	6	AX020429
12	838	28.8	838	6	AX009482
13	814.8	28.0	10719	1	AE006043
14	775.2	26.6	1893	1	AF064791
15	744.2	25.6	11545	1	U32772
16	744.2	25.6	110000	6	AR274543_09
17	433.8	14.9	300521	1	AE017153
18	410.8	14.1	873	6	AR377276
19	352.2	12.1	304250	1	EX294452
20	329.4	11.3	10687	1	AE015577
21	314.8	10.8	304558	1	AE017151
22	292.2	10.0	807	6	AR139598
23	280.8	9.7	10474	1	AE004472
24	280.6	9.6	10029	1	AE013918
25	280.6	9.6	208050	1	AE141414
26	278.2	9.6	349960	1	EX571861
27	278.2	9.4	10242	1	AX770908
28	273	9.4	26658	1	AE005511
29	265.8	9.1	891	6	AR377359
30	264.2	9.1	305325	1	AE016765
31	261.4	9.0	310613	1	AE016875
32	260.2	8.9	10405	1	AE000366
33	260.2	8.9	12943	1	AE015298
34	260.2	8.9	292906	1	AE016968
35	259.6	8.9	71128	1	ECU29581
36	258	8.9	1473	1	STYLGI
37	258	8.9	21583	1	AE008838
38	256.8	8.8	879	6	AR386236
39	256.4	8.8	230050	1	AL627277
40	256.4	8.8	301311	1	AE016843
41	248.2	8.5	876	6	AX188961
42	248.2	8.5	1310	1	ECOLGT
43	248.2	8.3	301086	1	AE016964
44	240.8	8.0	303438	1	AE017212
45	232.6	8.0	303438	1	AE017212

ALIGNMENTS

RESULT 1
VCH6514
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

VCH6514
Vibrio cholerae lgt and thya genes.
AF006514
AF006514.1 GI:3201563
lgt gene, proliptoprotein diacylglycerol transferase; thya;
thymidylate synthetase.
Vibrio cholerae
Bacteriophage; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.

2909 bp DNA linear BCT 05-JUN-1998
Carlén, N.I.A., Nilsson, A., Todorovic, M., Holmgren, J., and Lebens, M.
Characterisation of the lgt/thya locus from Vibrio cholerae

[REDACTED]

Db	6	TTGGATCGGCGCCCTAGAGCGGTGGCGCTGGTAAGGCTTATGATATTGGTGGGTTCTCTTT	1.20
Qy	121	TGCTATGTGTTGGCCCAATCGCCAGCGGCAATCGCGCGGCAGTGTGACCGGTGACA	1.80
Db	121	TGCTATGTGTTGGCCCAATCGCCAGCGGCAATCGCGCGGCAGTGTGACCGGTGACA	1.80
Qy	181	AGTCTGTGACTGTTATTGCGCGGCTTTTAAAGTGTAGTATCGGTGGCCGAGTTGTTA	2.40
Db	181	AGTCTGTGACTGTTATTGCGCGGCTTTTAAAGTGTAGTATCGGTGGCCGAGTTGTTA	2.40
Qy	241	TGTGATCTTCACAAATTTTGAATCTGTTCCTTGAACCCCTTTATTTATTCAAAGTGTG	3.00
Db	241	TGTGATCTTCACAAATTTTGAATCTGTTCCTTGAACCCCTTTATTTATTCAAAGTGTG	3.00
Qy	301	GATCGCGGCAGTCTCTTCACGCGGCTTATTGGGTGTATCAACGCACATGTTCTGTGA	3.60
Db	301	GATCGCGGCAGTCTCTTCACGCGGCTTATTGGGTGTATCAACGCACATGTTCTGTGA	3.60
Qy	361	TGCGCGTAAAAACCAACGACCTTCTTTGGTGTGGCCGAAATTTGTTGCCCTTTAGTGCC	4.20
Db	361	TGCGCGTAAAAACCAACGACCTTCTTTGGTGTGGCCGAAATTTGTTGCCCTTTAGTGCC	4.20
Qy	421	ATTGCGTTTGGGAGTAGGAGCGTATCGGAACTTTATCAATAGTGAATTGTGGGAGAGAT	4.80
Db	421	ATTGCGTTTGGGAGTAGGAGCGTATCGGAACTTTATCAATAGTGAATTGTGGGAGAGAT	4.80
Qy	481	AAGGATGTGCTTGGGCTTTGTATTTCTTATCTGTAATGTGCGCCATCGCGCGCATCTTTC	5.40
Db	481	AAGGATGTGCGCTTGGGCTTTGTATTTCTTATCTGTAATGTGCGCCATCGCGCGCATCTTTC	5.40
Qy	541	ACAGCTTATAGAAATCGCCTTAGAAGCGGTGTCTTCTTATTTCTTAATTTGGTTAT	6.00
Db	541	ACAGCTTATAGAAATCGCCTTAGAAGCGGTGTCTTCTTATTTCTTAATTTGGTTAT	6.00
Qy	601	TGGTAAACCTGTCGCGCTAGGACAGCGTATCGGACGTCTTTATAGCCTGTATAGGTAAT	6.60
Db	601	TGGTAAACCTGTCGCGCTAGGACAGCGTATCGGACGTCTTTATAGCCTGTATAGGTAAT	6.60
Qy	661	CCGCTTCTTGTGAAATACGTCCGTGAGCAGATGCTCAATTGGGTCTGTGTGGTGGCTT	7.20
Db	661	CCGCTTCTCTTGTGAAATACGTCCGTGAGCAGATGCTCAATTGGGTCTGTGTGGTGGCTT	7.20
Qy	721	CATTTCATGGGGCAATCCCTCTCCTTACCTAATAGTATATCGGTAATTTGATATGAT	7.80
Db	721	CATTTCATGGGGCAATCCCTCTCCTTACCTAATAGTATATCGGTAATTTGATATGAT	7.80
Qy	781	TTGGTCTTTCACAGCGCGGTTGTATTCAGAACCGGTGTAGCAGCAAAATAGGTAATTTAGT	8.40
Db	781	TTGGTCTTTCACAGCGCGGTTGTATTCAGAACCGGTGTAGCAGCAAAATAGGTAATTTAGT	8.40
Qy	841	GAACAAGTATTAGATCTTTGTCAAGCCGATCGTGCATCAAGGTGTTGGGTTGAAATAGA	9.00
Db	841	GAACAAGTATTAGATCTTTGTCAAGCCGATCGTGCATCAAGGTGTTGGGTTGAAATAGA	9.00
Qy	901	ACGAACGGGCAAGCGTTGTGAATGTGATTAATGCGAATTTGACCTACAGATGTGGGAA	9.60
Db	901	ACGAACGGGCAAGCGTTGTGAATGTGATTAATGCGAATTTGACCTACAGATGTGGGAA	9.60
Qy	961	CAATCAGTTTCTCTAGTGACTACACGAAGAATTTTGAAGAGTGCCTGATGCGGAGTT	10.20
Db	961	CAATCAGTTTCTCTAGTGACTACACGAAGAATTTTGAAGAGTGCCTGATGCGGAGTT	10.20
Qy	1021	GCTCGGTAATTCGTGGTTACGATATAGCGGCGGATTTTGCACATTAAGTATCCAAAC	10.80
Db	1021	GCTCGGTAATTCGTGGTTACGATATAGCGGCGGATTTTGCACATTAAGTATCCAAAC	10.80
Qy	1081	CTGGGATGCTTAATGCAATTTAAACAGCATGACCTTACCTTACCTGTAAGAGTGA	11.40
Db	1081	CTGGGATGCTTAATGCAATTTAAACAGCATGACCTTACCTTACCTGTAAGAGTGA	11.40
Qy	1141	GGATGACATGGGACGCGTGTATGGGTTCAGGCTAGAGCTTGGGCTTAAGCTGATGTGGT	12.00
Db	1141	GGATGACATGGGACGCGTGTATGGGTTCAGGCTAGAGCTTGGGCTTAAGCTGATGTGGT	12.00

QY 1201 TCATATTGACCACTTGAAAAAGATTGTGATGTTTGAAGCGCTGGCGTTGATGACCGAG 1260
 Db 1201 TCATATTGACCACTTGAAAAAGATTGTGATGTTTGAAGCGCGGTGATGATGACCGAG 1260
 QY 1261 TGAATTTCTTAACCTTCTACATCCGGGTGAAATTTCAATGAGGGTGTGTTGGCGCTTGCA 1320
 Db 1261 TGAATTTCTTAACCTTCTACATCCGGGTGAAATTTCAATGAGGGTGTGTTGGCGCTTGCA 1320
 QY 1321 GTACAGCCATCATTTTTCATGCTGGGGGATACCTTGATCTTCAACAGTATCAAGCTTC 1380
 Db 1321 GTACAGCCATCATTTTTCATGCTGGGGGATACCTTGATCTTCAACAGTATCAAGCTTC 1380
 QY 1381 ATGTGATGTCCTTGGGGGTGAATTTCAACATGATGTCAGGTTTATGTCCTTGCGCT 1440
 Db 1381 ATGTGATGTCCTTGGGGGTGAATTTCAACATGATGTCAGGTTTATGTCCTTGCGCT 1440
 QY 1441 GATGCAACAGATCACAGGAAAAAGCCGGGCTTGCGGTATCAACAGATGTCAGTCCGA 1500
 Db 1441 GATGCAACAGATCACAGGAAAAAGCCGGGCTTGCGGTATCAACAGATGTCAGTCCGA 1500
 QY 1501 CATTTACCAAGATCACTCGAATGATGCGCGATGTCAGCTTAAACGTGACCATTC 1560
 Db 1501 CATTTACCAAGATCACTCGAATGATGCGCGATGTCAGCTTAAACGTGACCATTC 1560
 QY 1561 AGCGCTCACTTCATATCATCAATCAAGATTAACACTGACAGATTTGAAATTTGGT 1620
 Db 1561 AGCGCTCACTTCATATCATCAATCAAGATTAACACTGACAGATTTGAAATTTGGT 1620
 QY 1621 CACTTGTGATGATTTGAGTCAACCGGATATCACTTCAAGATCTTATTCATACCT 1680
 Db 1621 CACTTGTGATGATTTGAGTCAACCGGATATCACTTCAAGATCTTATTCATACCT 1680
 QY 1681 TTCACTATCCCGTATTCAGCGGATGCTGCTGATGCTTTTATATTAATAAAGCTCC 1740
 Db 1681 TTCACTATCCCGTATTCAGCGGATGCTGCTGATGCTTTTATATTAATAAAGCTCC 1740
 QY 1741 CGAAGTGGGAGCTTTTATATCAAGATGATGCTTAAACGCTTAAAGCGTTAGGCGAG 1800
 Db 1741 CGAAGTGGGAGCTTTTATATCAAGATGATGCTTAAACGCTTAAAGCGTTAGGCGAG 1800
 QY 1801 AATGCTGCGGGGATGACGAAACACCGCAATTAAGTAACTCACCACCACTTTGCT 1860
 Db 1801 AATGCTGCGGGGATGACGAAACACCGCAATTAAGTAACTCACCACCACTTTGCT 1860
 QY 1861 CTTACAGCCCAAGTATGATGCTCAGACCTTTATATAGGCACTTCGCTAAGAAAG 1920
 Db 1861 CTTACAGCCCAAGTATGATGCTCAGACCTTTATATAGGCACTTCGCTAAGAAAG 1920
 QY 1921 AATACCGTAAATCAAGACCGTACCTCAAGTTAAAGCTTAAAGTGCACCGCAATTTG 1980
 Db 1921 AATACCGTAAATCAAGACCGTACCTCAAGTTAAAGCTTAAAGTGCACCGCAATTTG 1980
 QY 1981 CAGAGCAACACGCGAACTCAACAGACAGCGGTTGCGGAGAGAGAGAGTAAACA 2040
 Db 1981 CAGAGCAACACGCGAACTCAACAGACAGCGGTTGCGGAGAGAGAGAGTAAACA 2040
 QY 2041 AGTCCCAATGTTGCGACCTTAAGGTAATGGTATGATTTACGCACTTTACAGCGCCAGA 2100
 Db 2041 AGTCCCAATGTTGCGACCTTAAGGTAATGGTATGATTTACGCACTTTACAGCGCCAGA 2100
 QY 2101 GCCCAAGAGGAAACATTAAGGCTGTTGCTGATGATGATTAACCTAAACCGGTAC 2160
 Db 2101 GCCCAAGAGGAAACATTAAGGCTGTTGCTGATGATGATTAACCTAAACCGGTAC 2160
 QY 2161 CACTGACTGAAGCAATACCGTATGATGCTGCGGCAATGCAATTTTGTAGATTTTC 2220
 Db 2161 CACTGACTGAAGCAATACCGTATGATGCTGCGGCAATGCAATTTTGTAGATTTTC 2220
 QY 2221 AGGTGGGGGCAACCATTAACCTTCACTGATGTTTGGCACTACCGTTAATGCGACGA 2280
 Db 2221 AGGTGGGGGCAACCATTAACCTTCACTGATGTTTGGCACTACCGTTAATGCGACGA 2280

QY 2281 AATGTCGCAATTAACCAATACGATTAAGTGCACACACCGCAAGTATTAACCAATACCGA 2340
 Db 2281 AATGTCGCAATTAACCAATACGATTAAGTGCACACACCGCAAGTATTAACCAATACCGA 2340
 QY 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGTATCGTTGATTAATTAATCAAAAC 2400
 Db 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGTATCGTTGATTAATTAATCAAAAC 2400
 QY 2401 TTTCATGCTCATATGCGCAGTGCAGAGAGGCGAAACGAGCCAGTGTAGACTTTCTC 2460
 Db 2401 TTTCATGCTCATATGCGCAGTGCAGAGAGGCGAAACGAGCCAGTGTAGACTTTCTC 2460
 QY 2461 TAAATGCAACATCAATTTCTAGAGTGAAGAAATACAGACCGCGGAAATTAATAA 2520
 Db 2461 TAAATGCAACATCAATTTCTAGAGTGAAGAAATACAGACCGCGGAAATTAATAA 2520
 QY 2521 ATCGTGAATGCTGACCTGCGGAAAGACCGCGAAACTCTTTCTTAAGCGCATATGCGC 2580
 Db 2521 ATCGTGAATGCTGACCTGCGGAAAGACCGCGAAACTCTTTCTTAAGCGCATATGCGC 2580
 QY 2581 AAGCTGACGAGATTTGCTGACAGATGATCAATATTTGCAACCATACCATAGGAT 2640
 Db 2581 AAGCTGACGAGATTTGCTGACAGATGATCAATATTTGCAACCATACCATAGGAT 2640
 QY 2641 CGCGTTTCAACCGGTAACCCACCGCAACGAGCAACAAATATATAGAGTCAACGCTGCT 2700
 Db 2641 CGCGTTTCAACCGGTAACCCACCGCAACGAGCAACAAATATATAGAGTCAACGCTGCT 2700
 QY 2701 TGAGGATTAATGATGCTGCGGTTGCGCACTAAACCAATCATCATCTGCAATGGGTGGA 2760
 Db 2701 TGAGGATTAATGATGCTGCGGTTGCGCACTAAACCAATCATCATCTGCAATGGGTGGA 2760
 QY 2761 AGCAATTAATGAATGATTTGCTGATGCGCGGTTGCCATTAACCGCTGCCGAC 2820
 Db 2761 AGCAATTAATGAATGATTTGCTGATGCGCGGTTGCCATTAACCGCTGCCGAC 2820
 QY 2821 CATGCGACTGCAAGAAAGTAAATACAGATGAAGAACCAAGTTGCCAAGCTAGGCC 2880
 Db 2821 CATGCGACTGCAAGAAAGTAAATACAGATGAAGAACCAAGTTGCCAAGCTAGGCC 2880
 QY 2881 TTTGCTGCTCAGCGAATTCGCGCTGCGAG 2909
 Db 2881 TTTGCTGCTCAGCGAATTCGCGCTGCGAG 2909

RESULT 2
 AX009481
 LOCUS 2909 bp DNA linear PAT 06-SEP-2000
 DEFINITION Sequence 1 from Patent WO9961634.
 ACCESSION AX009481
 VERSION AX009481.1 GI:9996766
 KEYWORDS

SOURCE
 ORGANISM
 Vibrio cholerae
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.

REFERENCE
 1
 Carlin,N. and Lebens,M.R.
 Method of producing thy a<->strains of vibrio cholerae, such
 strains and their use
 Patent: WO 9961634-A 1 02-DEC-1999;
 JOURNAL CARLIN NLS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
 FEATRES
 Location/Qualifiers
 1..2909
 /organism="Vibrio cholerae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:666"

ORIGIN

Query Match 100.0%; Score 2909; DB 6; Length 2909;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGAGAGTTGTTATATGCTTACAGGGTATCTGCACTTTCACAAATATGACCCCGATATGTT	60
DB	1	GAGAGAGTTGTTATATGCTTACAGGGTATCTGCACTTTCACAAATATGACCCCGATATGTT	60
QY	61	TTGCATGAGCCCTCTACCGGTGCGCTGATGCGCTTATTTGGTGGGTTTCCTTTT	120
DB	61	TTGCATGAGCCCTCTACCGGTGCGCTGATGCGCTTATTTGGTGGGTTTCCTTTT	120
QY	121	TGCTATGTGTTGGCCATATGCGGAGGGATATCGCGCGGAGATGGTTGAGCGGGTAGCA	180
DB	121	TGCTATGTGTTGGCCATATGCGGAGGGATATCGCGCGGAGATGGTTGAGCGGGTAGCA	180
QY	181	AGTCTCTGACTGTATATTCGCCGCTTTTAAAGTATGATCGGTGCGCGAGTTGGTTA	240
DB	181	AGTCTCTGACTGTATATTCGCCGCTTTTAAAGTATGATCGGTGCGCGAGTTGGTTA	240
QY	241	TGTGATCTTCTACAAATTTGATCTGTTCCTTCTGACACCTCTTATTTATTCAAAGTGTG	300
DB	241	TGTGATCTTCTACAAATTTGATCTGTTCCTTCTGACACCTCTTATTTATTCAAAGTGTG	300
QY	301	GACTGGGGGAGATGTCCTTCCACCGCGGCTTATTTGGTGTGATCACCGCATGTTCTGTGA	360
DB	301	GACTGGGGGAGATGTCCTTCCACCGCGGCTTATTTGGTGTGATCACCGCATGTTCTGTGA	360
QY	361	TGCGCGTAAAAACCAACGACCTTCTTTGTGTGCGCGATTTGTTGCCCTTTATGTGCC	420
DB	361	TGCGCGTAAAAACCAACGACCTTCTTTGTGTGCGCGATTTGTTGCCCTTTATGTGCC	420
QY	421	ATTGCGTTGGGGATGAGACGTATACGGTAACTTTATGAATATGATCAACCGCATGTTCTGTGA	480
DB	421	ATTGCGTTGGGGATGAGACGTATACGGTAACTTTATGAATATGATCAACCGCATGTTCTGTGA	480
QY	481	AACGATGTGCTTGGGCTTTTGTATTCCTTAATGATGCGCCACTGCGCGCATCTTTC	540
DB	481	AACGATGTGCTTGGGCTTTTGTATTCCTTAATGATGCGCCACTGCGCGCATCTTTC	540
QY	541	ACAGCTTATTAATTCGCTTAGAAGGCGTGTCTGTCTTATTTCTTAATTTGGTTAT	600
DB	541	ACAGCTTATTAATTCGCTTAGAAGGCGTGTCTGTCTTATTTCTTAATTTGGTTAT	600
QY	601	TGCTAAACCTGCTCGCTAGGACAGAGTACCGGACTGTTTATAGCTGATACGTCATTT	660
DB	601	TGCTAAACCTGCTCGCTAGGACAGAGTACCGGACTGTTTATAGCTGATACGTCATTT	660
QY	661	CCGCTTCCTTGTGAATATGTCGCTGACCGACGATGCTCATGTTGGTCTGTTTGGTGGCTT	720
DB	661	CCGCTTCCTTGTGAATATGTCGCTGACCGACGATGCTCATGTTGGTCTGTTTGGTGGCTT	720
QY	721	CATTTCATGGGGCAAAATCTCTCTCTTACCTATGGTATCATTCGGTATTTTGTATGATGAT	780
DB	721	CATTTCATGGGGCAAAATCTCTCTCTTACCTATGGTATCATTCGGTATTTTGTATGATGAT	780
QY	781	TTGTCTTACAAACCGGCTTGTATCAAGACCTGTATGACAGACAAATATGGGTAGTTAGGT	840
DB	781	TTGTCTTACAAACCGGCTTGTATCAAGACCTGTATGACAGACAAATATGGGTAGTTAGGT	840
QY	841	GAAACAGATTTAAGTATCTTTGACAGCGGATGCTGATCAAGTGTTCGGTTGAAATATGA	900
DB	841	GAAACAGATTTAAGTATCTTTGACAGCGGATGCTGATCAAGTGTTCGGTTGAAATATGA	900
QY	901	ACGAACGGGACAGCGCTTGTGACTGTGATTAATGCGCATTTGACTTACATGATGTGGGCAA	960
DB	901	ACGAACGGGACAGCGCTTGTGACTGTGATTAATGCGCATTTGACTTACATGATGTGGGCAA	960
QY	961	CAATCAGTTTCTCTTAGTACTTACACGCAAGATTTTGGAAAGCTGCGCATGCGAGTT	1020
DB	961	CAATCAGTTTCTCTTAGTACTTACACGCAAGATTTTGGAAAGCTGCGCATGCGAGTT	1020
QY	1021	GCTCGGCTAATTCGTGTGATGAGATATGCGGCGGATTTTGGCCATTTAGGTACCAAAAC	1080
DB	1021	GCTCGGCTAATTCGTGTGATGAGATATGCGGCGGATTTTGGCCATTTAGGTACCAAAAC	1080
QY	1081	CTGGATGCTAATGCGCAATTTAACCAGATGCTCAACATCTTACGTTAAAGTATGA	1140

Db	1081	CTGGGATGCTAATGCCAATTTTAACCAAGCATGGCTCAACAACTCTTACCGTAAGGTGA	1140
Oy	1141	GGATGACATGGGAGCGCGTATGTGTTCAAGGTAGACCTTGGCTTAAGCTGATGTGG	1200
Db	1141	GGATGACATGGGAGCGCGTATGTGTTCAAGGTAGACCTTGGGCTTAAGCTGATGTGG	1200
Oy	1201	TCTATTATGACCACTTGAAGAAAGTTGTATGATTTGACCCGGCGGGCTTGATGACCGAGG	1260
Db	1201	TCTATTATGACCACTTGAAGAAAGTTGTGATGATTTGAGCCGTGGCGCTTGATGACCGAGG	1260
Oy	1261	TGAATCTCTTAACCTTGACATCCGGGTGTAATTTCAATGGGTGTGTTGGCCCTTGCAAT	1320
Db	1261	TGAATCTCTTAACCTTGACATCCGGGTGTAATTTCAATGGGTGTGTTGGCCCTTGCAAT	1320
Oy	1321	GTAACGCCATCAATTTTTCATTGCTGGGGGATACCTTGATCTCTCAADAGTACAGGGTTC	1380
Db	1321	GTAACGCCATCAATTTTTCATTGCTGGGGGATACCTTGATCTCTCAADAGTACAGGGTTC	1380
Oy	1381	ATGGATGTGCCCCCTGGGGGTTGTAATTTCAACATGTGCGAGGTTATGTGTCTTGCGCT	1440
Db	1381	ATGGATGTGCCCCCTGGGGGTTGTAATTTCAACATGTGCGAGGTTATGTGTCTTGCGCT	1440
Oy	1441	GATGGCACAGATCAACAGGAAAAAGCCGGGCTTGGCGTATCAACAATCGTCAATGGCA	1500
Db	1441	GATGGCACAGATCAACAGGAAAAAGCCGGGCTTGGCGTATCAACAAGTCGTCAATGGCA	1500
Oy	1501	CATTACCAAGATCAACTGGAATTGATGGCGGATGTGCAGCTTAAACGTAGAGCATTTCC	1560
Db	1501	CATTACCAAGATCAACTGGAATTGATGGCGGATGTGCAGCTTAAACGTAGAGCATTTCC	1560
Oy	1561	AGCGCCTCAGTTCCATATCAATCCAAAGATTAAAACTGCGAGGATTTGGAAACTTGGGT	1620
Db	1561	AGCGCCTCAGTTCCATATCAATCCAAAGATTAAAACTGCGAGGATTTGGAAACTTGGGT	1620
Oy	1621	CACTTTGGATGATTTTAAAGCTCAACCCGGATATCAATTCAGTTCACAGTCTTATTAACCGGT	1680
Db	1621	CACTTTGGATGATTTTAAAGCTCAACCCGGATATCAATTCAGTTCACAGTCTTATTAACCGGT	1680
Oy	1681	TTTCAGTCTAAATCCCGTATTCAGGCGGTATGGCTTGATGGGTTTATATAAAAAAGCTCC	1740
Db	1681	TTTCAGTCTAAATCCCGTATTCAGGCGGTATGGCTTGATGGGTTTATATAAAAAAGCTCC	1740
Oy	1741	CGAAGGTGGGAGCTTTTTTTTATACATGATGCTTTAAAGCTTAACCGGTTTGGGGAAG	1800
Db	1741	CGAAGGTGGGAGCTTTTTTTTATACATGATGCTTTAAAGCTTAACCGGTTTGGGGAAG	1800
Oy	1801	AATGCTGCGGGGATGACCAACACACCAACCAATTAATCTACACCACTTTTGCT	1860
Db	1801	AATGCTGCGGGGATGACCAACACACCAACCAATTAATCTACACCACTTTTGCT	1860
Oy	1861	CTTACAAGCCCAAGTTGATGAGCTTAGACCTTTTATATGACGAGTTGGGTAGGAAG	1920
Db	1861	CTTACAAGCCCAAGTTGATGAGCTTAGACCTTTTATATGACGAGTTGGGTAGGAAG	1920
Oy	1921	AATACCGTAATTCAAAGCCGTAGGCATCAAGTTAAAGCTTAAGTGACACAGCGCAATTTG	1980
Db	1921	AATACCGTAATTCAAAGCCGTAGGCATCAAGTTAAAGCTTAAGTGACACAGCGCAATTTG	1980
Oy	1981	CAGAGCAAAACCGGCAAACTCACCAAGACACGCGTTGGCGGACAGACGATTAATCA	2040
Db	1981	CAGAGCAAAACCGGCAAACTCACCAAGACACGCGTTGGCGGACAGAGAGTAAATCA	2040
Oy	2041	AGTGGCAATGTTTGCACCTAAGATTAAGTTAGGTTATTCACGCACTTTACAGACGCCAGA	2100
Db	2041	AGTGGCAATGTTTGCACCTAAGATTAAGTTAGGTTATTCACGCACTTTACAGACGCCAGA	2100
Oy	2101	GCCACAGAGAGAAACATTAAGCTGTGTGTGTCGATGAAGATTGAATTAATACCGTAAC	2160
Db	2101	GCCACAGAGAGAAACATTAAGCTGTGTGTGTCGATGAAGATTGAATTAATACCGTAAC	2160
Oy	2161	CACGTACCTGAAGCAATACCGTGTAGTGGGCTTCGGGCATTCGCATTTTGTGAATTTCC	2220

Db 2161 CACTGTACTGAAGCAATACCGTGTAGTGGGCTCGGCCAATCGCATTGTAAGATTTC 2220

QY 2221 ACCGCGCGGCGCAACCACTCAAACTCTTCACTCACTTTGCCCATACCGTAATGGCGAGAA 2280

Db 2221 ACCGCGCGGCGCAACCACTCAAACTCTTCACTCACTTTGCCCATACCGTAATGGCGAGAA 2280

QY 2281 AATGTCGCAATACCCCAATACGATTAAGTGCAGACCCAGCAAGATTAATCCCAATCCGA 2340

Db 2281 AATGTCGCAATACCCCAATACGATTAAGTGCAGACCCAGCAAGATTAATCCCAATCCGA 2340

QY 2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGTATGTTGATTAATCAAAACC 2400

Db 2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGTATGTTGATTAATCAAAACC 2400

QY 2401 TTTCATCTCATATGCGCAGTCGCGAAGCAGAGCGGCAAGCGCAGTGGAGCTTTC 2460

Db 2401 TTTCATCTCATATGCGCAGTCGCGAAGCAGAGCGGCAAGCGCAGTGGAGCTTTC 2460

QY 2461 TAAATATCCCAACATCATTTCTAGAGGTAGAAAGATCAGACCGCGAGAAATTTGAAAA 2520

Db 2461 TAAATATCCCAACATCATTTCTAGAGGTAGAAAGATCAGACCGCGAGAAATTTGAAAA 2520

QY 2521 ATGTGTATGTGGGCTGGCGGAAAGCAGCGGCAAACTCTTCTTACAGCGCATATGGCC 2580

Db 2521 ATGTGTATGTGGGCTGGCGGAAAGCAGCGGCAAACTCTTCTTACAGCGCATATGGCC 2580

QY 2581 AAGGCTGACGAGAGATTTGCTCAGATGATACCAATTTGGCACCACCATAGAGAT 2640

Db 2581 AAGGCTGACGAGAGATTTGCTCAGATGATACCAATTTGGCACCACCATAGAGAT 2640

QY 2641 CGGCGTTTCAACCGGTATACCCACCGGCAAGAGACCAACATATATGAAGTCACTGCT 2700

Db 2641 CGGCGTTTCAACCGGTATACCCACCGGCAAGAGACCAACATATATGAAGTCACTGCT 2700

QY 2701 TGAAGATTTGATCAGTCCGCTTGCCACTAAACCATCATCATCTGCAATTTGGGAGGA 2760

Db 2701 TGAAGATTTGATCAGTCCGCTTGCCACTAAACCATCATCATCTGCAATTTGGGAGGA 2760

QY 2761 AGCAAAATTCAAATAGAACTTTGCTGTATGCGCGGTTGGCCATTAAAAACCGCTGGCAG 2820

Db 2761 AGCAAAATTCAAATAGAACTTTGCTGTATGCGCGGTTGGCCATTAAAAACCGCTGGCAG 2820

QY 2821 CATCGGAGCTGCAAGAGATTAATATACGATGAAAGCCAGATTGGCCCAAGTAGGCC 2880

Db 2821 CATCGGAGCTGCAAGAGATTAATATACGATGAAAGCCAGATTGGCCCAAGTAGGCC 2880

QY 2881 TTTCGTGTGTCAGGCAATCGCGCTGCGAG 2909

Db 2881 TTTCGTGTGTCAGGCAATCGCGCTGCGAG 2909

RESULT 3
AB004153
LOCUS
DEFINITION
Vibrio cholerae O1 biovar eltor str. N16961 chromosome I, section 61 of 251 of the complete chromosome.
ACCESSION
AE004153 AE003852
VERSION
AE004153.1 GI:9655103
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae O1 biovar eltor str. N16961
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
REFERENCE
AUTHORS
Heidelberg, J.F., Eissen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bess, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D., Nieman, W.C., and White, O.
TITLE
DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae
JOURNAL
Nature 406 (6795), 477-483 (2000)

MEDLINE
PUBMED
REFERENCE
AUTHORS
20406833
10952301
2 (bases 1 to 9931)
Heidelberg, J.F., Eissen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bess, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D., Nieman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
TITLE
JOURNAL
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
FEATURES
source
1. 9931
/organism="Vibrio cholerae O1 biovar eltor str. N16961"
/mol_type="genomic DNA"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:243277"
/chromosome="1"
/note="biotype: El Tor"
288. 602
/gene="VC0666"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP93831.1"
/db_xref="GI:9655104"
/translation="VSSDFSLIVIPSKTTGENTMRSIYLOLVLAIVLAEREERVMKRVRRSSHYLPMNNPHLRDILETDGPRIGLSLPAVVAERVRHRRVLAATPT"
complement (640. 1698)
/gene="VC0667"
complement (640. 1698)
/note="similar to PID:882727 GB:U00096 PID:1789199 PID:2342573; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="oxidoreductase Tas, aldo/keto reductase family"
/protein_id="AAP93832.1"
/db_xref="GI:9655105"
/translation="MALRGSAMQVTKLPHSSLEISKICGTMFGHNSQAAPQOLDYALERGVNFTIDREMYVPPTROTQKTEFIGNMLAKGKREKIVLATKVGPRVPIYDKRALDHRNTHQAVDSDSLRLQTDYTDQLQHWFORQTFPGQINPYDKQEEVTILETLALNDLVRMGKRVYIGVSNETPWGWSYLRARKEHLPRTVSIQNYNLINRSFEVGLAEISHEGVGLAVSPPLAFGLNGKYLNGARPAAGACTLHCRSRYFTEGILAEVAVLAQOFGDLPQAQALAFVQRPFAVSNITGATTMEOQLKSNLSDISLVAELLOKIOEIGTITSNCP"
complement (1891. 2556)
/gene="VC0668"
complement (1891. 2556)
/note="similar to GB:Y00113 SP:P06722 GB:U16361 PID:42065 PID:577425; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="DNA mismatch repair protein MucH"
/protein_id="AAP93833.1"
/db_xref="GI:9655106"
/translation="MKRAPYTOQLITRAQIAGLSFAELADBAQTVPPDLKXGMVGOILLEHLATAGSRPQOPFHLGIEIKSIPISTYGGPLETFVCAVPLGVGHGLTWGSHVNRKSKVIMPIVQGRERIPLARCGSPGLMSPEEBAOLKADBEEMENIVLGWVAQITAKHGSVQLRPPAANGRLTETVANGNRPKRLPRGFIYLRQFTQIIORVYA"
complement (2943. 3035)
/gene="VC0669"
complement (2943. 3035)
/gene="VC0669"

```

/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFP9383.1"
/db_xref="GI:9655107"
/translation="MRHTAHLFTFSVNNRVKKGCGVKGKGIYQ"
3018..3164
/gene="VC0670"
3018..3164
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFP9383.1"
/db_xref="GI:9655108"
/translation="MRGSHKIKVDNLQNLSPFIPFTNKYRANHSKAPLGDWVK
NORN"
3124..3705
/gene="VC0671"
3124..3705
/note="similar to PID:882723 GB:U00096 SP:Q46930
PID:1789194; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="murt/nudix family protein"
/protein_id="AAFP9383.1"
/db_xref="GI:9655109"
/translation="MLRIGICGKINAKFEKTRGLPVIDGGYELNAGIVICNNHGVF
WAKRYGDSWQFQGGIDGDSPEQAMRELYEVGLTKDVIVATSHHMYKLPK
RLVWDSQVQICQKQKFLRLCECDSEKINMQGSSPEFDGWRWVSYWYPRQVVSF
KRDYRRMRKEFASLAMPFERKVKYGRNTHRG"
3708..5954
/gene="VC0672"
3708..5954
/note="similar to PID:882722 GB:U00096 PID:1255724
PID:1789193; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoenolpyruvate-protein phosphotransferase"
/protein_id="AAFP9383.1"
/db_xref="GI:9655110"
/translation="WLSQKRDIVQVSKEDVYQALDIPKQTCAMSTECTVILAN
BEHRLELMATQGLKFKGDKIHIDFEGVLGVSAPLNTLAASKHPSFKYFQGLG
BEVHSLGTPIIYKQVGLVAVQKSPRLFSEMEVSPVLTAAQLAVLVAHQTOG
HMRLSKQQAITGVAASGVALGEFTWDDTQPDLSALDPFEQERLALAVEN
ALADPRRRKLDGIDINKALAIEDLFTHLNDPMCKDKQKQIQGDRADALQOV
ESYNNRFPMSDYLRRADQIRELIGRLIFLLNTHNPOKIDKRVILVVELTASV
LAALPKXILAVSLRGAANSHAAIISRAIGTPAWGVSINPDLNGKKGIVDYSK
LFTSPKTIINERYALANERBELSRMNEAIRPACTLGGARIELILNAGLSADTSA
VNGVDGGLRTRISFLDHRFPSEBEQYQYRHVLANYPORVVMRFLDGGKPL
PYLPREDNPLGWRGIRFTLDHPIFIQLRAMLASASBGLSILLPMISGAELD
DAKFIYQAYQEVQODPRVMPQIGIMEVPSMLILDLKIDPVSAGNDLTOY
LLAVDRNNARADYVESMHPAVMALQIQCTCAQIPEVCQGLAGDPICALLIG
LGTTLSMNTSNVAKYKYLVRHSELAEITQALQALTOPEGSEIYVMMLAYIEKGFPA
GEVRAGKK"
5964..6758
/gene="VC0673"
5964..6758
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFP9383.1"
/db_xref="GI:9655111"
/translation="WNTIELAILLLIGAVVGLAGLIGGGLLVVPAALLFLPYAGI
STELVQALATSLATILITSGSSALNHLKGNVDVAVVKKMPGVGGFSGVAV
MISQYLPKFGVIVLCTISVQWRSRALQNHMPASAMITLTCGIGIIVISLAIQV
GSLVPLNHRHGIEMKKAIGSSVCGFALAIQSMIGFILHGVQVNTLPOYSIGVYLP

```

```

ALAIATTSMLTRIGAKLATQMPARLKRFPALFIMCVAVTMLPQ"
/gene="VC0674"
6787..7602
/gene="VC0674"
6787..7602
/note="similar to GP:3201564; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="prolipoprotein diacylglycerol transferase"
/protein_id="AAFP9383.1"
/db_xref="GI:9655112"
/translation="WPGYIOPFNIDPVPFSTGPLAVRYGLMILVGLFANMLANR
ADRGSGWTRQVSDLLFAGFLGVVIGSVGVYIFYNEDLFLADPLYLEKWTGMSF
HGDLYGVTIMFWYARKNORTPEGVADFPAPVPLQMGRLGNMNSLMDVDP
WAFVFPNGPLPHPSQLVEFALGVLEFLLNMFPGKRPPIGSGVGLFAGYTFRF
LVEYREPDQALQLFGFISMGQILSLPMVITIGILMMVSVYKRGYQDRAAK"
7612..8463
/gene="VC0675"
7612..8463
/note="similar to GP:3114981; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thymidylate synthase"
/protein_id="AAFP9384.1"
/db_xref="GI:9655113"
/translation="MRQYDLDCRIYDQGVYENERTGRCLITVINALTYDVGNQF
PLVTRKSPKAAVABLLGYIRGYDAAADFRLGRTKRDANNLQAMLNRYRGED
DMGRVYGVQKRAVAKPDGSHIDQKKIYDLRGVYDGEIINFNPEFPMGLRPLC
MYHHNSLIDGLTYLNSRSCDVPGLNFMVQVYFALAAQITGKPGIAYKIV
NAHYOQOELMADVOLKEPFPAPQFINPKIKLQDLETWTLDDVDVGYQHPD
IQYFQSV"
Query Match 98.7%; Score 2870.6; DB 1; Length 9911;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2885; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
1 GAGAGGTTGTATGCTCAGGTTATCTGAGTTCCCAATATGACCCCGATTGTT 60
6774 GAGAGGTTGTATGCTCAGGTTATCTGAGTTCCCAATATGACCCCGATTGTT 6833
61 TTGATGCGCCCTTACCGTGCGCTGATAGCTTGAATTTTGGTGGTTCCTTT 120
6634 TTGATGCGCCCTTACCGTGCGCTGATAGCTTGAATTTTGGTGGTTCCTTT 6893
121 TTGATGCGCCCTTACCGTGCGCTGATAGCTTGAATTTTGGTGGTTCCTTT 180
6894 TTGATGCGCCCTTACCGTGCGCTGATAGCTTGAATTTTGGTGGTTCCTTT 6953
181 AGTCTGATGTTGATATGCGCGGCTTTTAAAGTGAATCGGTGGCCGAGTTGTTA 240
6954 AGTCTGATGTTGATATGCGCGGCTTTTAAAGTGAATCGGTGGCCGAGTTGTTA 7013
241 TGTGATCTTACAAATTTGATGTTGCTTGTCTGACCTCTTATTTATCAAGTGTG 300
7014 TGTGATCTTACAAATTTGATGTTGCTTGTCTGACCTCTTATTTATCAAGTGTG 7073
301 GACTGGCGGATGCTCTTACGCGGCGCTTATGGGTGATACACCGCATGTTGGTGA 360
7074 GACTGGCGGATGCTCTTACGCGGCGCTTATGGGTGATACACCGCATGTTGGTGA 7133
361 TGGCGGTAACCAACGACCTTCTTGTGTGCGCATTTTGTGCGCCCTTATGAGTCC 420
7134 TGGCGGTAACCAACGACCTTCTTGTGTGCGCATTTTGTGCGCCCTTATGAGTCC 7193
421 ATTGCGTTGGGATGGACGATAGCTTATGAAATGAACTTTGGGACGAGT 480
7194 ATTGCGTTGGGATGGACGATAGCTTATGAAATGAACTTTGGGACGAGT 7253
481 AAGGATGTCCTTGGGCTTTTGTATTCCTAATGATGAGCCACGCGCGCATCTTC 540

```

Db 7254 AACGATGTCCTGGGCTTTGATTCCTAATGATGAGCCACACTGCCGCGCATCTTC 7313
 Qy 541 AACGCTTATGATTTGGCTTTAGAAAGCGTGGTTCCTTATTTCTTAATGTTAT 600
 Db 7314 AACGCTTATGATTTGGCTTTAGAAAGCGTGGTTCCTTATTTCTTAATGTTAT 7373
 Qy 601 TGGTAAACCTGTCGCTAGAGCAGCGATCCGAGCTGTTTATGCTGATAGCTAT 660
 Db 7374 TGGTAAACCTGTCGCTAGAGCAGCGATCCGAGCTGTTTATGCTGATAGCTAT 7433
 Qy 661 CCGCTTCTTGTGAAATACGTCGCTAGAGCAGCGATCCGAGCTGTTTATGCTGAT 720
 Db 7434 CCGCTTCTTGTGAAATACGTCGCTAGAGCAGCGATCCGAGCTGTTTATGCTGAT 7493
 Qy 721 CATTTAAATGGGGCAATTCCTCCCTTACCTATGATGATCATGCTGATTTGATGAT 780
 Db 7494 CATTTAAATGGGGCAATTCCTCCCTTACCTATGATGATCATGCTGATTTGATGAT 7553
 Qy 781 TTGGTCTTACAAAGCGGCTTTGATCAAGACCGTGTAGAGCAAAATAGGTAAGT 840
 Db 7554 TTGGTCTTACAAAGCGGCTTTGATCAAGACCGTGTAGAGCAAAATAGGTAAGT 7613
 Qy 841 GAAACGATTTATGATCTTTGTGAGCGCATCGTCATCAAGGTGTTGGGTTGAAATGA 900
 Db 7614 GAAACGATTTATGATCTTTGTGAGCGCATCGTCATCAAGGTGTTGGGTTGAAATGA 7673
 Qy 901 ACGAAGCGGCAAGCGTGTGATGCTGTGATTAATGCGATTTGACCTACGATGGGCA 960
 Db 7674 ACGAAGCGGCAAGCGTGTGATGCTGTGATTAATGCGATTTGACCTACGATGGGCA 7733
 Qy 961 CAATCATGTTTCTGTAGTACATACGCAAGAGTTTGGAAAGCTGCCGTAAGCGAGT 1020
 Db 7734 CAATCATGTTTCTGTAGTACATACGCAAGAGTTTGGAAAGCTGCCGTAAGCGAGT 7793
 Qy 1021 GCTCGGCTATATTCGCTGATGATGATTAATGCGGCGATTTTCCGCAATTTAGTACCAAAAC 1080
 Db 7794 GCTCGGCTATATTCGCTGATGATGATTAATGCGGCGATTTTCCGCAATTTAGTACCAAAAC 7853
 Qy 1081 CTGGGATGCTAATGCCATTTAAACCAAGCATGCTCAACATCTTACCGTAAAGTGA 1140
 Db 7854 CTGGGATGCTAATGCCATTTAAACCAAGCATGCTCAACATCTTACCGTAAAGTGA 7913
 Qy 1141 GATGATCATGGAAGCGTGTATGCTGTGAGGTTAGAGCTTGGGCTTAAGCTGATGTGG 1200
 Db 7914 GATGATCATGGAAGCGTGTATGCTGTGAGGTTAGAGCTTGGGCTTAAGCTGATGTGG 7973
 Qy 1201 TCATATGACCAAGTGAATAAAGATTGATGATTTGAGCCGTGGCGTTGATGACGAG 1260
 Db 7974 TCATATGACCAAGTGAATAAAGATTGATGATTTGAGCCGTGGCGTTGATGACGAG 8033
 Qy 1261 TGAATATCTTAATCTTCAATCCGCGTGAATTTCAATGCGGCTTTGGCGCTTGCAT 1320
 Db 8034 TGAATATCTTAATCTTCAATCCGCGTGAATTTCAATGCGGCTTTGGCGCTTGCAT 8093
 Qy 1321 GTACAGCCATCATTTTTCATGCTGGGGATACCTGATTCATCAACATGATCAGCTTC 1380
 Db 8094 GTACAGCCATCATTTTTCATGCTGGGGATACCTGATTCATCAACATGATCAGCTTC 8153
 Qy 1381 ATGTATGTGCTTGGGCTTGAATTTCAATGCTGAGCTTTATGCTTCTTGGCT 1440
 Db 8154 ATGTATGTGCTTGGGCTTGAATTTCAATGCTGAGCTTTATGCTTCTTGGCT 8213
 Qy 1441 GATGACAGATCAACAGGAAAGCGCGGCTTGGCGTATCAACATGCTGATGCGCA 1500
 Db 8214 GATGACAGATCAACAGGAAAGCGCGGCTTGGCGTATCAACATGCTGATGCGCA 8273
 Qy 1501 CATTTACCAAGATCAATCGAATGATGCGGATGAGCTTAAACGCTGAGCAATTTCC 1560
 Db 8274 CATTTACCAAGATCAATCGAATGATGCGGATGAGCTTAAACGCTGAGCAATTTCC 8333
 Qy 1561 AGCGGCTCAGTTTCCATTCATCCAAAGATTAAACATGCGAGATTGGAACTTGGT 1620
 Db 8334 AGCGGCTCAGTTTCCATTCATCCAAAGATTAAACATGCGAGATTGGAACTTGGT 8393

Qy 1621 CACTTGTGATGATTTGACGTCACCGGATATCAGTTCACAGATCTATTCATAACCGTT 1680
 Db 8394 CACTTGTGATGATTTGACGTCACCGGATATCAGTTCACAGATCTATTCATAACCGTT 8453
 Qy 1681 TTGAGCTAATCCGCTATTCAGGCGGCTATGCTGATGCTGTTTATATATATATATAT 1740
 Db 8454 TTGAGCTAATCCGCTATTCAGGCGGCTATGCTGATGCTGTTTATATATATATATAT 8513
 Qy 1741 CGAAGTGGGAGCTTTTATATACAGATGATGCTTTAAGCCTTAAAGCGGTTAGGCAAG 1800
 Db 8514 CGAAGTGGGAGCTTTTATATACAGATGATGCTTTAAGCCTTAAAGCGGTTAGGCAAG 8573
 Qy 1801 AATGCTGCGGGGATGACGACAAACACCAATPAAGTACTACCAACCAATTTGCT 1860
 Db 8574 AATGCTGCGGGGATGACGACAAACACCAATPAAGTACTACCAACCAATTTGCT 8633
 Qy 1861 CTTACAAACCCCAAGTGTGATGATGAGCTCAGACCTTAAATGCGGCTTGGGTAAGG 1920
 Db 8634 CTTACAAACCCCAAGTGTGATGATGAGCTCAGACCTTAAATGCGGCTTGGGTAAGG 8693
 Qy 1921 AATACCTGTAATCAAGACCGTATGACCATCAAGTTAAAGCTTAAAGTCAACAGCGCATTTG 1980
 Db 8694 AATACCTGTAATCAAGACCGTATGACCATCAAGTTAAAGCTTAAAGTCAACAGCGCATTTG 8753
 Qy 1981 CAGAGCAAAACGCGCAAACTCACAGAGACAGCGGCTGGCGGCGAGAGCATATTAACA 2040
 Db 8754 CAGAGCAAAACGCGCAAACTCACAGAGACAGCGGCTGGCGGCGAGAGCATATTAACA 8813
 Qy 2041 AGTCCATATGTTCCACCTAAGTAAGTAATGATTTCAAGCCTTTCAGACGCGCAGA 2100
 Db 8814 AGTCCATATGTTCCACCTAAGTAATGATTTCAAGCCTTTCAGACGCGCAGA 8873
 Qy 2101 GCCCAGAGAGAACCATTTAGCTGTTGTGTGATGAAAGTTGAATTAATACGTAAC 2160
 Db 8874 GCCCAGAGAGAACCATTTAGCTGTTGTGTGATGAAAGTTGAATTAATACGTAAC 8933
 Qy 2161 CACTGTACCTGAAGCAATACCGTGTAGTGGGCGCTGGCGCAATGCAATTTGTGAATTC 2220
 Db 8934 CACTGTACCTGAAGCAATACCGTGTAGTGGGCGCTGGCGCAATGCAATTTGTGAATTC 8993
 Qy 2221 AAGTGGCGGCAACATCAAACTCTTCATCAAGTTTCCCATCAACCGTAATGGCGAGA 2280
 Db 8994 AAGTGGCGGCAACATCAAACTCTTCATCAAGTTTCCCATCAACCGTAATGGCGAGA 9053
 Qy 2281 AATGTCGCAATCCCAATAGCATTAAGTGGAGACCAACCGAAGTATACCAATCCGA 2340
 Db 9054 AATGTCGCAATCCCAATAGCATTAAGTGGAGACCAACCGAAGTATACCAATCCGA 9113
 Qy 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTGTGATATGCTTTGTAATAACCAAAAC 2400
 Db 9114 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTGTGATATGCTTTGTAATAACCAAAAC 9173
 Qy 2401 TTTCAATGCTCATATCCGCAATCCGAGAGAGCGAAACGACGATGTGAGCTTCTC 2460
 Db 9174 TTTCAATGCTCATATCCGCAATCCGAGAGAGCGAAACGACGATGTGAGCTTCTC 9233
 Qy 2461 TAAATGCAAAATCATTTCTAGAGTGAAGATACAGACCGCGGAAAGATTGAAAAA 2520
 Db 9234 TAAATGCAAAATCATTTCTAGAGTGAAGATACAGACCGCGGAAAGATTGAAAAA 9293
 Qy 2521 ATGTGATGTGTGACATGCGCAAGACGCGCAAACTCTTCTTAAACGCGCATATGCGC 2580
 Db 9294 ATGTGATGTGTGACATGCGCAAGACGCGCAAACTCTTCTTAAACGCGCATATGCGC 9353
 Qy 2581 AAGGCTGACAGAGTATGCTGACAGTATGACCAATTTGGACCGCATCACTAAGAAAT 2640
 Db 9354 AAGGCTGACAGAGTATGCTGACAGTATGACCAATTTGGACCGCATCACTAAGAAAT 9413
 Qy 2641 CGCGGTTTCAACCGGTAACCCACCGGCAACGAGACCAACATTAATGAAGTCACTGTCT 2700
 Db 9414 CGCGGTTTCAACCGGTAACCCACCGGCAACGAGACCAACATTAATGAAGTCACTGTCT 9473

QY 2701 TGAGGATTGAATCACTGCGCTTGGCACTAAACCAATCATCATCTGCAATTGGGATGGGA 2760
DB 9474 TAGAGATTGAATCACTGCGCTTGGCACTAAACCAATCATCATCTGCAATTGGGATGGGA 9533
QY 2761 AGCAATTCGAATGAATCTTGGCTGATGCGCGCTTGGCACTAAACCAATGGCGGAC 2820
DB 9534 AGCAATTCGAATGAATCTTGGCTGATGCGCGCTTGGCACTAAACCAATGGCGGAC 9593
QY 2821 CATCGGACCTGCAAGAACTAGTAATAACAGCATGAAAGCCAGTTTGGCCCAAGTACGCC 2880
DB 9594 CATCGGACCTGCAAGAACTAGTAATAACAGCATGAAAGCCAGTTTGGCCCAAGTACGCC 9653
QY 2881 TTTGGTGTGACGCAAGAACTGCGCTGCGAG 2909
DB 9654 TTTGGTGTGACGCAAGAACTGCGCTGCGAG 9682

RESULT 4
VCTHYAG 1410 bp DNA linear BCT 01-NOV-2000
LOCUS Vibrio cholerae thya gene.
DEFINITION Y17135
VERSION Y17135.1 GI:3114980
KEYWORDS thya gene; thymidilate synthase.
SOURCE Vibrio cholerae
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.

REFERENCE 1
AUTHORS Valle, E., Ledon, T., Cedre, B., Campos, J., Valmaseda, T., Rodriguez, B., Garcia, L., Marrero, K., Benitez, J., Rodriguez, S. and Fando, R.
TITLE Construction and characterization of a nonproliferative El Tor cholera vaccine candidate derived from strain 638
JOURNAL Infect. Immun. 68 (11), 6411-6418 (2000)
MEDLINE 20490577
PUBMED 11035753
REFERENCE 2 (bases 1 to 1410)
AUTHORS Benitez, J.A.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1998) J.A. Benitez, Centro Nacional de Investigaciones, Cientificas, PO Box 6990, La Habana, CUBA

FEATURES
source
1..1410
/organism="Vibrio cholerae"
/mol_type="genomic DNA"
/strain="C7258"
/db_xref="taxon:666"
289..1140
/gene="thya"
289..1140
/gene="thya"
/codon_start=1
/transf_table=11
/product="thymidilate synthase"
/protein_id="CAA645.1"
/db_xref="GI:3114981"
/db_xref="GOA:O66108"
/db_xref="SWISS-PROT:O66108"
/translation="MROYLDCRIVDQGWENERTGKCLTVINADLTVDVGNNOPLVITTRKSPKAAVAELIAYIRGNDNAADPRLGCTMTMANINLNCAPRYNPRGCDPMGRVYGVGRAMAKPDGHIIDILKIVDDLSGVDREIILNPNRPHYRGICRPMYSHHSFLGDTIPLNLSRSCDPEGLNFMVQYVPLALPAQITGKPGELAKTIVNAHYDQDELMDVOKREPFPAQFPHINPKITLQDLEWTLDLDPDVGXGFHPDIQYFVSF"

ORIGIN
Query Match 47.8%; Score 1390.8; DB 1; Length 1410;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1398; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 551 GAATTCGCTTGAAGAGCGGTCTTCTTATTTCTTAATGTTTATGTAACCT 610
|||||
DB 1 GAATTCGCTTGAAGAGCGGTCTTCTTATTTCTTAATGTTTATGTAACCT 610

DB 1 GAATTCGCTTGAAGAGCGGTCTTCTTATTTCTTAATGTTTATGTAACCT 610
QY 611 CGTCGCTAGACAGCGGATCCGACCTGTTTATAGTGATACGATACATTCGCTTCCT 670
DB 61 CGTCGCTAGACAGCGGATCCGACCTGTTTATAGTGATACGATACATTCGCTTCCT 120
QY 671 GTGGAATAGCTCCGTGAGCCAGATGCTCACTTGGGTCTGTTGGCTTCACTTCAATG 730
DB 121 GTGGAATAGCTCCGTGAGCCAGATGCTCACTTGGGTCTGTTGGGTCTTCAATG 180
QY 731 GGGCAATTCCTCTCTTACCTATGATGATCATGGAATTTGATAGTGTGCTTAC 790
DB 181 GGGCAATTCCTCTCTTACCTATGATGATCATGGAATTTGATAGTGTGCTTAC 240
QY 791 AAGCGCGTTTGTATCAAGACGCTGTGACAGCAAAATAGGGTATGTTAGTGAACAGTAT 850
DB 241 AAGCGCGTTTGTATCAAGACGCTGTGACAGCAAAATAGGGTATGTTAGTGAACAGTAT 300
QY 851 TTGATCTTTGTCAGGCAATCGTCATCAAGGTGTTGGTGAATAATGAACGAGGAC 910
DB 301 TTGATCTTTGTCAGGCAATCGTCATCAAGGTGTTGGTGAATAATGAACGAGGAC 360
QY 911 AAGCGTTTGTATCTGTGATTAATGCCATTTGACCTACAGATGAGGCAACATAGTTT 970
DB 361 AAGCGTTTGTATCTGTGATTAATGCCATTTGACCTACAGATGAGGCAACATAGTTT 420
QY 971 CCTCTAGTACTACAGCAAGATTTTGAAGCTGCGGTAGCGAGTTGGTGGCTAT 1030
DB 421 CCTCTAGTACTACAGCAAGATTTTGAAGCTGCGGTAGCGAGTTGGTGGCTAT 480
QY 1031 ATTGCTGTATGATTAATGCGGCGATTTTGGCCATTAAGTACCAAACTGGGATGCT 1090
DB 481 ATTGCTGTATGATTAATGCGGCGATTTTGGCCATTAAGTACCAAACTGGGATGCT 540
QY 1091 AATGCCAATTTAAACCAAGATGCTGCAACATCTCTCCGTAAGGATGAGATGATGATG 1150
DB 541 AATGCCAATTTAAACCAAGATGCTGCAACATCTCTCCGTAAGGATGAGATGATGATG 600
QY 1151 GAGCGGTGTATGCTGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 1210
DB 601 GAGCGGTGTATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 660
QY 1211 CAGTTGAAAAGATTTGTATGATGATTTGAGCCGCTGCGTGAACCGAGTGAATTTCT 1270
DB 661 CAGTTGAAAAGATTTGTATGATGATTTGAGCCGCTGCGTGAACCGAGTGAATTTCT 720
QY 1271 AACTTCAATCCGCGGTAAATTTCAATGCGGTGTTTGGCCTTGTGATGACCAT 1330
DB 721 AACTTCAATCCGCGGTAAATTTCAATGCGGTGTTTGGCCTTGTGATGACCAT 780
QY 1331 CATTTTCAATGCTGCGGTAAATTTCAATGCGGTGTTTGGCCTTGTGATGATG 1390
DB 781 CATTTTCAATGCTGCGGTAAATTTCAATGCGGTGTTTGGCCTTGTGATGATG 840
QY 1391 CCTTGGGCTGATTTCAATGCTGCAAGGTGATGCTGCTGCTGCTGCTGCTGCTG 1450
DB 841 CCTTGGGCTGATTTCAATGCTGCAAGGTGATGCTGCTGCTGCTGCTGCTGCTG 900
QY 1451 ATCAGAGGGAAGAGCGGCTTGGCGTATCAAGATGCTGCAATGCGCATTTTACCA 1510
DB 901 ATCAGAGGGAAGAGCGGCTTGGCGTATCAAGATGCTGCAATGCGCATTTTACCA 960
QY 1511 GATCACTGCAATGATGCTGCGATGCTGCAAGGTGATGCTGCTGCTGCTGCTGCTG 1570
DB 961 GATCACTGCAATGATGCTGCGATGCTGCAAGGTGATGCTGCTGCTGCTGCTGCTG 1020
QY 1571 TTCATATCAATCAAGATTTAAACACGTCAGATTTGAAACTGGGTCACTTTGGAT 1630
DB 1021 TTCATATCAATCAAGATTTAAACACGTCAGATTTGAAACTGGGTCACTTTGGAT 1080
QY 1631 GATTTGACGTACCGGATATCATGTTCAACATCTTCAATACCCGTTTCACTTAA 1690
DB 1081 GATTTGACGTACCGGATATCATGTTCAACATCTTCAATACCCGTTTCACTTAA 1140

QY 1691 TCCCGTATTCAGCGGGTATGGCTTGATGCGTTTATATATAAAAGTCCCGAAGTCCG 1750
DB 1141 TCCCGTATTCAGCGGGTATGGCTTGATGCGTTTATATATAAAAGTCCCGAAGTCCG 1200
QY 1751 GAGCTTTTATATACAGATGATGCTTTAAGCTTTAAGCGGTAGGCGCAAGATGTCGCCG 1810
DB 1201 GAGCTTTTATATACAGATGATGCTTTAAGCTTTAAGCGGTAGGCGCAAGATGTCGCCG 1260
QY 1811 GGGATGACGCAAAACACACCCCAATAGTAATCAACACCACTTTGCTTTCAAGCC 1870
DB 1261 GGGATGACGCAAAACACACCCCAATAGTAATCAACACCACTTTGCTTTCAAGCC 1320
QY 1871 CAAGTTGAGATGAGCTCAGACCTTTAATAGGCAAGTTCGGTAAAGGAAGATACCGTAA 1930
DB 1321 CAAGTTGAGATGAGCTCAGACCTTTAATAGGCAAGTTCGGTAAAGGAAGATACCGTAA 1380
QY 1931 ATCAAGACCGGTAGCCATCAAGTTAAAGCTT 1960
DB 1381 ATCAAGACCGGTAGCCATCAAGTTAAAGCTT 1410

RESULT 5
AX009483 1222 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 3 from Patent WO961634.
ACCESSION AX009483
VERSION AX009483.1 GI:9996768
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae
Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.

REFERENCE
1
AUTHORS Carlin, N. and Lebens, M.R.
TITLE Method of producing thy a<->-strains of vibrio cholerae, such
as strains and their use
JOURNAL Patent: WO 961634-A 3 02-DEC-1999;
CARLIN NILS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
FEATURES
source
1. .1222
/organism="Vibrio cholerae"
/mol_type="unassigned DNA"
/db_xref="taxon:686"

ORIGIN
Query Match 42.0%; Score 1222; DB 6; Length 1222;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1688 TAAATCCCGTATTCAGCGGGTATGGCTTGATGCGTTTATATATAAAAGTCCCGAAGT 1747
DB 1 TAAATCCCGTATTCAGCGGGTATGGCTTGATGCGTTTATATATAAAAGTCCCGAAGT 60
QY 1748 CCGGAGCTTTTATACAGATGATGCTTTAAGCTTTAAGCGGTAGGCGCAAGATGCTG 1807
DB 61 CCGGAGCTTTTATACAGATGATGCTTTAAGCTTTAAGCGGTAGGCGCAAGATGCTG 120
QY 1808 CCGGAGTACGACAAACACACCCCAATAGTAATCAACACCACTTTGCTTTCA 1867
DB 121 CCGGAGTACGACAAACACACCCCAATAGTAATCAACACCACTTTGCTTTCA 180
QY 1868 GCCCAAGTTGAGATGAGCTCAGACCTTTAATAGGCAAGTTCGGTAAAGGAAGATACCG 1927
DB 181 GCCCAAGTTGAGATGAGCTCAGACCTTTAATAGGCAAGTTCGGTAAAGGAAGATACCG 240
QY 1928 TAAATCAAGACCGGTAGCCATCAAGTTAAAGCTTTAAGTACACCGCAATTTGCGAGCA 1987
DB 241 TAAATCAAGACCGGTAGCCATCAAGTTAAAGCTTTAAGTACACCGCAATTTGCGAGCA 300
QY 1988 AACACGGCAAACTCAACCGAGACGACGGTGGGCGAGACAGACGATATCAAGTGCCA 2047
DB 301 AACACGGCAAACTCAACCGAGACGACGGTGGGCGAGACAGACGATATCAAGTGCCA 360

QY 2048 ATGTTGCAACCTTAAGTAATAGGTAGTATTTCAACGCACTTTCAAGCAGCCAGAGCCACG 2107
DB 361 ATGTTGCAACCTTAAGTAATAGGTAGTATTTCAACGCACTTTCAAGCAGCCAGAGCCACG 420
QY 2108 AGAGGAACCTTAAGCTGCTGTTGCTGATGAAGTTAATCAATTAATCCGTAAACCTGTA 2167
DB 421 AGAGGAACCTTAAGCTGCTGTTGCTGATGAAGTTAATCAATTAATCCGTAAACCTGTA 480
QY 2168 CCGTGAAGCAATACCGTGTAGTGGGCTGGCCCAATGCGATTTTGTAGATTTCAAGTGGC 2227
DB 481 CCGTGAAGCAATACCGTGTAGTGGGCTGGCCCAATGCGATTTTGTAGATTTCAAGTGGC 540
QY 2228 CCGGCAACCAATCAACTTTTATCATGATTTGCCCATACCGTAAATGCGCAAGAAATGTC 2287
DB 541 CCGGCAACCAATCAACTTTTATCATGATTTGCCCATACCGTAAATGCGCAAGAAATGTC 600
QY 2288 GCAATATCCCAATACGATTAAGTGCACACCGCAAGATTAACCAATACCGAAAGCTGG 2347
DB 601 GCAATATCCCAATACGATTAAGTGCACACCGCAAGATTAACCAATACCGAAAGCTGG 660
QY 2348 GTTTCAGCCCTGTGATGACAGGTTTGTGTAATCGGTTTGAATTAATCAAACTTTTCATG 2407
DB 661 GTTTCAGCCCTGTGATGACAGGTTTGTGTAATCGGTTTGAATTAATCAAACTTTTCATG 720
QY 2408 CTCATATGCGCAGTCCGACAGAGGCGGAAACGAGCCAGTGTGACTTTCTTAAATG 2467
DB 721 CTCATATGCGCAGTCCGACAGAGGCGGAAACGAGCCAGTGTGACTTTCTTAAATG 780
QY 2468 CCAACATCATTTCTTAAGTGTAGAGTAAGTAATGACACCGCGCAAGATTAAGAAATTCGNG 2527
DB 781 CCAACATCATTTCTTAAGTGTAGAGTAAGTAATGACACCGCGCAAGATTAAGAAATTCGNG 840
QY 2528 ATGTGGGACCTGCGGAAAGCAACCGCGCAACTCTTTTAAAGCGCATATGCGCAAGCTG 2587
DB 841 ATGTGGGACCTGCGGAAAGCAACCGCGCAACTCTTTTAAAGCGCATATGCGCAAGCTG 900
QY 2588 ACGAGATATGTCACAGTAATGATTAATGTCACCCATGACCATAGAGATGCGGCT 2647
DB 901 ACGAGATATGTCACAGTAATGATTAATGTCACCCATGACCATAGAGATGCGGCT 960
QY 2648 TCAACCGGTAAACCAACCGGCAACGACCAACCAATTAATGATGACCGTGTGAGAT 2707
DB 961 TCAACCGGTAAACCAACCGGCAACGACCAACCAATTAATGATGACCGTGTGAGAT 1020
QY 2708 TGAATCAGTCCGCTGTCACATTAACCAATCATTAATCTTGTGCTGGGAGCAAT 2767
DB 1021 TGAATCAGTCCGCTGTCACATTAACCAATCATTAATCTTGTGCTGGGAGCAAT 1080
QY 2768 TCAATTAAGAACTTTGCTGATGCGCGGTGCGCATTTAAACCGTGGCGCAACATGCG 2827
DB 1081 TCAATTAAGAACTTTGCTGATGCGCGGTGCGCATTTAAACCGTGGCGCAACATGCG 1140
QY 2828 ACTGCAAGATGTAGTAATACAGATTAAGCAAGTTTCCCAAGCTTGTGCTG 2887
DB 1141 ACTGCAAGATGTAGTAATACAGATTAAGCAAGTTTCCCAAGCTTGTGCTG 1200
QY 2888 GTCAGCGAAATCGGCGTGCAG 2909
DB 1201 GTCAGCGAAATCGGCGTGCAG 1222

RESULT 6
AY143429 1236 bp DNA linear BCT 28-SEP-2002
LOCUS Vibrio cholerae mutant thymidilate synthetase (thyA) gene, complete
DEFINITION cds.
ACCESSION AY143429
VERSION AY143429.1 GI:23343945
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae
Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

REFERENCE 1 (bases 1 to 1236)
 AUTHORS Xia X.
 TITLE Direct Submission
 JOURNAL Submitted (20-AUG-2002) Microbiology, Institute for Infectious Disease Control and Prevention, Chinese Center for Disease Control and Prevention, P.O. Box 5, Changping, Beijing 102206, P.R. China

FEATURES
 source
 1..1236
 /organism="Vibrio cholerae"
 /mol_type="genomic DNA"
 /db_xref="taxon:666"

gene

CDS

248..1099
 /gene="thyA"
 /note="thyA"
 /gene="thyA"
 /note="does not catalyze the conversion of dUTP into dTMP"
 /codon_start=1
 /transl_table=11
 /product="mutant thymidilate synthetase"
 /protein_id="AA016904.1"
 /db_xref="GI:23343946"
 /translation="MRQYLDLCRIVDQGVWENERTGRCLVINADLYDVGNNG
 PLVTRKSPFMAVAELGYIRGYNADPQLGTGTWDANANILNPNPKRGE
 DMGRVYGVGRAMAKPDGGHIDOLKIVDLSRGVDSEILNFPNPEFNGCLAPC
 MYSHPSLIGDLYLNSRSCDPLMKNMNVQYVFLAIAOITGKPGIAKIV
 NHHYDDLEMDVQLKREPPAPQPHINPKIKLQDLTMTWLTDDSDVGYQPHD
 IQYPSV"

ORIGIN

Query Match 41.8%; Score 1216.8; DB 1; Length 1236;
 Query Similarity 99.0%; Pred. No. 0;
 Matches 1224; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 592 TTGGTTATGTGTAACCTGTCCTAGGACGCGTATCCGACATGTTTATAGCTGATGA 651
 DB 1 TTGGTTATGTGTAACCTGTCCTAGGACGCGTATCCGACATGTTTATAGCTGATGA 60
 QY 652 CGGTACATTCGGCTTCCTGTGGAATACGTCGAGCAGATGCTCAGTTGGGCTGTT 711
 DB 61 CGGTACATTCGGCTTCCTGTGGAATACGTCGAGCAGATGCTCAGTTGGGCTGTT 120
 QY 712 TGGTGGCTTCATTCATGAGGCGCAATCCTCTCTACCTATGATGATCATGATATTT 771
 DB 121 TGGTGGCTTCATTCATGAGGCGCAATCCTCTCTACCTATGATGATCATGATATTT 180
 QY 772 GATGATGCTTGGCTTCATGAGGCGGCTTGTATCAAGCCGTGAGCAAGCAAAATAGGG 831
 DB 181 GATGATGCTTGGCTTCATGAGGCGGCTTGTATCAAGCCGTGAGCAAGCAAAATAGGG 240
 QY 832 TAGTTAGTGAACAGATATTAGATCTTGTGTCAGCGCATGTCGATCAAAGTGTGGGT 891
 DB 241 TAGTTAGTGAACAGATATTAGATCTTGTGTCAGCGCATGTCGATCAAAGTGTGGGT 300
 QY 892 TGAATAATGAACGAGCGGCAAGCGTGTGATGATGATTAAGCCGATTTGACCTAGCA 951
 DB 301 TGAATAATGAACGAGCGGCAAGCGTGTGATGATGATTAAGCCGATTTGACCTAGCA 360
 QY 952 TGTGGGCAACATAGTTTCCTCTAGTACATACGCAAGAGTTTGGAAAGCTCCGT 1011
 DB 361 TGTGGGCAACATAGTTTCCTCTAGTACATACGCAAGAGTTTGGAAAGCTCCGT 420
 QY 1012 AGCCGAGTGTGCTGCTATATTTGTGTTACGATATGCGCGGATTTTGGCAATTAAG 1071
 DB 421 AGCCGAGTGTGCTGCTATATTTGTGTTACGATATGCGCGGATTTTGGCAATTAAG 480
 QY 1072 TACCAAAACCTGGGATGCTATGCGCATTTAAACCAAGCATGGCTCAACATCCTTACCG 1131
 DB 481 TACCAAAACCTGGGATGCTATGCGCATTTAAACCAAGCATGGCTCAACATCCTTACCG 540
 QY 1132 TAAAGGTGAGATGACATGAGGACGCGTGTATGAGATTGAGGGTAGAGCTTGGGCTAAGCC 1191

DB 541 TAAAGGTGAGATGACATGAGGACGCGTGTATGAGATTGAGGGTAGAGCTTGGGCTAAGCC 600
 QY 1192 TGAATGCTGCTCATATTGACAGTGAAGAAAGATTGTTGATGATTTAGCCGTGCGCTTGA 1251
 DB 601 TGAATGCTGCTCATATTGACAGTGAAGAAAGATTGTTGATGATTTAGCCGTGCGCTTGA 660
 QY 1252 TGACCGAGGTGAATCTTAACTTCTTACCAATCCGGGTGAATTTACATGGGCTTTGGG 1311
 DB 661 TGACCGAGGTGAATCTTAACTTCTTACCAATCCGGGTGAATTTACATGGGCTTTGGG 720
 QY 1312 CCTTGGCATGTAAGGCATCATTTTTCATGCTGGGGGGATCTGTATCTCAACAGTAC 1371
 DB 721 CCTTGGCATGTAAGGCATCATTTTTCATGCTGGGGGGATCTGTATCTCAACAGTAC 780
 QY 1372 TCAGCGTTCATGATGTCGCCCTTGGGGTGAATTTCAACATGTCAGGCTTTATGCTTT 1431
 DB 781 TCAGCGTTCATGATGTCGCCCTTGGGGTGAATTTCAACATGTCAGGCTTTATGCTTT 840
 QY 1432 CCTTGGCATGATGTCGCCCTTGGGGTGAATTTCAACATGTCAGGCTTTATGCTTT 1491
 DB 841 CCTTGGCATGATGTCGCCCTTGGGGTGAATTTCAACATGTCAGGCTTTATGCTTT 900
 QY 1492 CATTGGGCACTTATACCAAGATCACTGATGATGATGTCGCGATGTCAGCTTAAACGTGA 1551
 DB 901 CATTGGGCACTTATACCAAGATCACTGATGATGATGTCGCGATGTCAGCTTAAACGTGA 960
 QY 1552 GCCATTCCAGCGCCTTCAGTTCATATCAATCCAAAGATTAAACCTGACAGATTTGGA 1611
 DB 961 GCCATTCCAGCGCCTTCAGTTCATATCAATCCAAAGATTAAACCTGACAGATTTGGA 1020
 QY 1612 AACTGGGTCACCTTGGATGATTTTGAACGTCACCGATATCAGTTCAGATCTTATTA 1671
 DB 1021 AACTGGGTCACCTTGGATGATTTTGAACGTCACCGATATCAGTTCAGATCTTATTA 1080
 QY 1672 ATACCCGTTTCACTGATTAATCCCGTATTCAGCGGGATGAGCTTGAATGCTTTATTA 1731
 DB 1081 ATACCCGTTTCACTGATTAATCCCGTATTCAGCGGGATGAGCTTGAATGCTTTATTA 1140
 QY 1732 AAAGCTCCGGAAGTGGGAGCTTTTATACAGATGATGCTTAAACGCTTAAAGCGGT 1791
 DB 1141 AAAGCTCCGGAAGTGGGAGCTTTTATACAGATGATGCTTAAACGCTTAAAGCGGT 1200
 QY 1792 TAGGGCAAGATGCTGCCGGGATGACGACAAACAC 1827
 DB 1201 TAGGGCAAGATGCTGCCGGGATGACGACAAACAC 1236

RESULT 7
 VCAJ10968/c
 LOCUS VCAJ10968 1591 bp DNA linear BCT 21-SEP-1998
 DEFINITION Vibrio cholerae nptA gene.
 ACCESSION AJ010968
 VERSION AJ010968.1 GI:3646476
 KEYWORDS nptA gene; phosphate pump.
 SOURCE Vibrio cholerae
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
 REFERENCE 1
 Lebens M., Soderlund L.O., Lundquist P. and Carlén N.I.A.
 A putative sodium dependent phosphate pump in Vibrio cholerae
 Unpublished
 JOURNAL 2 (bases 1 to 1591)
 REFERENCE
 Carlén N.I.A.
 Submitted (17-SEP-1998) Carlén N.I.A., SBL Vaccin AB, Department of Molecular Biology, PMB, UV 30, Stockholm, S-105 21, SWEDEN
 nptA gene is located between the nhar gene (AJ002395) and the
 thyA gene (AJ006514) both running in opposite directions.
 COMMENT
 FEATURES
 source
 1..1591
 /organism="Vibrio cholerae"

```
/mol_type="genomic DNA"
/strain="JS1569"
/db_xref="taxon:666"
245..1588
/gene="nplA"
/gene="nplA"
245..250
/gene="nplA"
267..272
/gene="nplA"
373..376
/gene="nplA"
383..1532
/gene="nplA"
/codon_start=1
/cran1_table=11
/product="sodium dependent phosphate pump"
/protein_id="CAA09443.1"
/db_xref="GI:3646477"
/db_xref="GOA:O87918"
/db_xref="SPTREMBL:O87918"
/translation="MINOATSPAAPISTLTGKLRWANLAFMLYLLLAAMVSGFK
ATGDAKLVFEFASHPIAGIMIGLVATLIGSSSTVTSIIIVLAVGGIPVETALPMV
GANIGTNTNIVSIGHWCKEPRAPASATIHDPNLLAVLIFELPMFGILEKV
SHWISPLIATGDMKSKFPIKPIITKPIITGLETOLSVLGTGVALIVGLATIF
VAITWGLMKSLMGRAREILQNAIGRPGLIASGTIVTVLVOSSTTSLMVLIV
GSGVLRREIYPTLGEANIGTITALLATATVGEFAVFAIQIALVHLSPMLMVLIV
YGIPLRELPIELISTWACKSKMNVVSYLVGVFVIVPSIILATLTA"
1550..1588
/gene="nplA"
/notes="shared with thyA gene running in opposite
direction"
/function="transcription termination"
```

stem_loop

ORIGIN

Query Match 40.8%; Score 1187; DB 1; Length 1591;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1723 TTTATTAATAAAAGCTCCCGAGAGTGGAGCTTTTATTAACAGATAGTCTTAAACG 1782
DB 1591 TTTATTAATAAAAGCTCCCGAGAGTGGAGCTTTTATTAACAGATAGTCTTAAACG 1532
QY 1783 TTAAGCGGTTAAGGCAAGATGCTGCGGGGATGACGCAAAACACCAACCAATTAAGTACT 1842
DB 1531 TTAAGCGGTTAAGGCAAGATGCTGCGGGGATGACGCAAAACACCAACCAATTAAGTACT 1472
QY 1843 CACCAACCAATTTGCTCTTCAAGCCCAAGTTAGATGAGTCAAGCCTTTAATAG 1902
DB 1471 CACCAACCAATTTGCTCTTCAAGCCCAAGTTAGATGAGTCAAGCCTTTAATAG 1412
QY 1903 CAGTGGGGTAAAGAAAGATTCCTTAATCAAGCCGTAGCCTCAAGTTAAAGCTTAA 1962
DB 1411 CAGTGGGGTAAAGAAAGATTCCTTAATCAAGCCGTAGCCTCAAGTTAAAGCTTAA 1352
QY 1963 GTGCAACGACGCAATTTGCAAGCAACGCAAACTCACACGAGACAGCGGTGGCG 2022
DB 1351 GTGCAACGACGCAATTTGCAAGCAACGCAAACTCACACGAGACAGCGGTGGCG 1292
QY 2023 GAGCAGACGAGTAATACAGTGCCTAATGTTGGACCTTAAGGTAAATGGTAGATTACG 2082
DB 1291 GAGCAGACGAGTAATACAGTGCCTAATGTTGGACCTTAAGGTAAATGGTAGATTACG 1232
QY 2083 CACTTTGACAGCGCAGACGCCACGAGAGAAACCACTTAGCTGGTGGTGGTGCATGAGA 2142
DB 1231 CACTTTGACAGCGCAGACGCCACGAGAGAAACCACTTAGCTGGTGGTGGTGCATGAGA 1172
QY 2143 TTGAATTAATACCGTAACCACTGTAACCTGTAACCAATCCGTGATGAGTGGCCCTCGCCAT 2202
DB 1171 TTGAATTAATACCGTAACCACTGTAACCTGTAACCAATCCGTGATGAGTGGCCCTCGCCAT 1112
QY 2203 CGCATTTTGTGAATTTTACAGGTGGGGCGGCAACCAATCAAACTCTTCACTGTTGGCCAT 2262
DB 1111 CGCATTTTGTGAATTTTACAGGTGGGGCGGCAACCAATCAAACTCTTCACTGTTGGCCAT 1052
```

QY	2263	CACCGTAATGCGAGCAAAATGCTGCAATACCAATACGTAATAGTGGACACCAAGCA	2322
DB	1051	CACCGTAATGCGAGCAAAATGCTGCAATACCAATACGTAATAGTGGACACCAAGCA	992
QY	2323	AGTATTACCAATACCGAAGACCTGGTTTCAAGCCCTGTGTGACAGTTTGTAAATCGG	2382
DB	991	AGTATTACCAATACCGAAGACCTGGTTTCAAGCCCTGTGTGACAGTTTGTAAATCGG	932
QY	2383	TTGATTAATAATCAAAACCTTCATGCTCATATCCGACGTCGCAAGACGAGCGAAGAG	2442
DB	931	TTGATTAATAATCAAAACCTTCATGCTCATATCCGACGTCGCAAGACGAGCGAAGAG	872
QY	2443	CCAGTGTAGACCTTCTCTTAATAAGCCAAACATATTTCTTAGAGTGTGAAGATCAGAC	2502
DB	871	CCAGTGTAGACCTTCTCTTAATAAGCCAAACATATTTCTTAGAGTGTGAAGATCAGAC	812
QY	2503	CGCAGAAAGATTGAATAATCGTGTGATGCTGCACTGGCGCAAGCAAGCGCAAACTCTTC	2562
DB	811	CGCAGAAAGATTGAATAATCGTGTGATGCTGCACTGGCGCAAGCAAGCGCAAACTCTTC	752
QY	2563	TTTACAGCGCATATGCGCAAGGCTGACGAGAGTATTGCTACAGTATACCAATATTGGC	2622
DB	751	TTTACAGCGCATATGCGCAAGGCTGACGAGAGTATTGCTACAGTATACCAATATTGGC	692
QY	2623	ACCATCACCATAGGAATCGGGTTTCAACCGGTAACCCACCGGCAACGAGACCAACAT	2682
DB	691	ACCATCACCATAGGAATCGGGTTTCAACCGGTAACCCACCGGCAACGAGACCAACAT	632
QY	2683	AATGAAGTACCGCTGCTTGAAGATTAATCAGTCCGTTGCCACTAAACCAATCATCA	2742
DB	631	AATGAAGTACCGCTGCTTGAAGATTAATCAGTCCGTTGCCACTAAACCAATCATCA	572
QY	2743	TCCGCAATTTGGGGGGAAGCAATTCATTAATCAATCTTGGCTTGAATCCCGGTTGCCA	2802
DB	571	TCCGCAATTTGGGGGGAAGCAATTCATTAATCAATCTTGGCTTGAATCCCGGTTGCCA	512
QY	2803	TTTAAACCGCTGCCGACCATTCGCGACTGCAAGAGTAAATTAATACATGAAAGCCAA	2862
DB	511	TTTAAACCGCTGCCGACCATTCGCGACTGCAAGAGTAAATTAATACATGAAAGCCAA	452
QY	2863	GTTTGCCCAACGTAGGCTTTCGTGTCAGCGAAATTCGCGCTGCAG 2903	
DB	451	GTTTGCCCAACGTAGGCTTTCGTGTCAGCGAAATTCGCGCTGCAG 405	
RESULT 8			
AP005074 298900 bp DNA linear BCT 05-MAR-2003			
LOCUS			
DEFINITION			
AP005074 BAO00031			
ACCESSION			
AP005074.1 GI:28805287			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Vibrio parahaemolyticus			
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;			
Vibrionaceae; Vibrio.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			

MEDLINE 22508454
 PUBMED 12620739
 REFERENCE 3 (bases 1 to 298900)
 AUTHORS Oshima, K., Kurokawa, K., Nakino, K., Yokoyama, K., Yasunaga, T.,
 Honda, T., Shinagawa, H., Hattori, M. and Iida, T.
 TITLE Direct Submission
 JOURNAL Submitted (09-APR-2002) Ken Kurokawa, Osaka University, Genome
 Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
 Japan (E-mail: kengen-info.osaka-u.ac.jp)
 URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365,
 Fax: 81-6-6879-2047)
 COMMENT This clone was isolated from a patient presenting with acute
 gastroenteritis.
 FEATURES
 source Location/Qualifiers
 1. 298900
 /organism="Vibrio parahaemolyticus"
 /mol_type="genomic DNA"
 /strain="O3:K6"
 /sub_strain="RIMD 2210633"
 /db_xref="taxon:670"
 /chromosome="1"
 115. 315
 /gene="VP0305"
 115. 315
 /note="similar to PIR:556442 percent identity 49 in 63 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="BAC58568.1"
 /db_xref="GI:28805288"
 /translation="MLREPAVVRPLQVAFVYKTLFPGQFFIVGVSFCEDNKKLLPD
 VKDQKRLAVREYNQALSLSA"
 complement (399. 1037)
 /gene="VP0306"
 complement (399. 1037)
 /note="similar to GB:AAF95690.1 (AE004324) percent
 identity 74 in 208 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="peptide methionine sulfoxide reductase"
 /protein_id="BAC58569.1"
 /db_xref="GI:28805289"
 /translation="MLNKOTLISIDALPGRBOPMOIEDCHFNOSLTAFLAHQOQ
 ILGMCGFCAERLPMQIDGVSIVSGVTPNPTEVEVCTGKTHGVYVYDE
 RVLSAQLAVFWEKDPDPCMGQNDLCTQVRSALYTSQDOELASKLOVQAL
 EAEIRSTITTEIVPAGPYFAETTHQOYLAKNPDGTCIGTGVCPPSLQG"
 1205. 2917
 /gene="VP0307"
 1205. 2917
 /note="similar to GB:AAF95689.1 (AE004323) percent
 identity 66 in 571 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="BAC58570.1"
 /db_xref="GI:28805290"
 /translation="MIKKALPIIVGLVVSQAVADVSLKLGIDALDNVAVLSS
 IPEKDVSTSRPQALDOSTEALNALGYHAKISYISSECDLIVNIHAGLPKIK
 IMDVIVSGEAKDEEPPNLAISPLVYGLINQERDSIKSIRNLALORGFYVSD
 VGEYNQSLSTWDFSSVVEPDLSKEDRELPKIVSLAPAKNQLETGIGSTDTGV
 RGTLMKKPWSARSHSFNTLSLRPEQTLTLAGYKPLDVLRYEYQQLGKMLDN
 RDTESLNSLAVERHWLTDGWHKTVYHLYENFSQGLDDGVQFVPLGAFSTRY
 RCGSMPEWGDQSVTEYVGDPLSETRVLLGRSSWIRGIGENHRLPREGGANT
 TEPEKSPSLRFAGGNNIRGYVESIPVDEGALGAKYILSTLEYRYVGN
 WMAATFDIDAFNDTPMKSGAGVIRWASPVGPSDFANGIDENKNERHFHSI

gene 2914. 6681
 /gene="VP0308"
 2914. 6681
 /note="similar to GB:AAF95688.1 (AE004323) percent
 identity 54 in 1258 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="BAC58571.1"
 /db_xref="GI:28805291"
 /translation="MIKKFVTKWLSLIGLLLLITIFVATVLFTHPGKALIMGA
 EKALPOLKVENIQGSLPFRFALONVSFDESILHIDAKYERIALAINFRCPDPKVCYD
 EALQGVNFWPELPPASHEPEETPLRISTPVPFVFNKSVENDIVANLQIDW
 QNSTLSWQGRVLVAPALKDINVLAPASAEBAEAEKEDTAPDQIVPEVW
 IPLTVARPLDTHNFKLAGETPVVILGIVARAGGVEDVKTTELMEVEGKLSQ
 VTSADPYPRKQDLALVQADKQKSLSSAGSVGDSLNNATSEIYQALIKNDIQP
 LKQLPDPIVYNVQAMPFQSDIYVSFSLAKGSLDGEYALERTKAGKQIPAY
 DVALLNGKGLLEQIDLESIVKTLGSELSEKVAWMAADINWQADLNLKNIQPGQWES
 AEGDISGSLSTSGSLTEQGVQVSLPKLIDIGILNGYPLNNGQLASDNKNGEDIQ
 LITQGLALSHGPNQLSAKGIDKQIMDEVEVFPDPFASVDPDLAGKQKVALGSLK
 EPDINDLANNQOQANVTITLKQDVTPLPAPKANLSLIANNITDYDIKVDSAD
 LEVSGDEKLHQLTDVSDVLSISLIGTKQKEMTMDGALRLRLTISQOGPMSIQ
 KSTAVKNTDKQIANVQACWLOAKSSVCLTEDISVGTQGAIXLAINPNDQIKOFL
 PPTTKVQSVMQAFKAKAPKKEPVTLSVNNPKQVQALAEQPIKNDPESRKLAL
 AKQKDAEWLPVKQNGDLSGKVSILNVSSEPTIDGVALSTHLDPLALICEYSL
 FXANNWTKLALSDVWHPKRVNQFLIDMKKQGEVTPIDINSQVINFKRQADVLA
 GILTPDGKLEITGDADVMPKQDMRTTRAPFKQELKVMPEVVKI KVEDMTIDTPOL
 AKVEGINLPMQRI VIBELPSAVGSSDVTIYNKQDLPVEVAMPEVNTDINIKI
 GDPOFLAFLGKGLKGLSINVTOKDKGPFIVGEVNIIVGSRSGQDVLIEGKILNN
 GPADQPYVSIKAIKRNPDNDQIVAGVATPASPQSVIEIPEEAMPANLSTLLG
 ODIDSSGQNAVTTLTILGSLAKGSRVAGELGEPAGQVODLOLDAGSGDSQVAVSGY
 ILPGQAVKVCIGTIRNSLSEFTVRLMDPLTLEAVSGVDSAVDLLIQEPRP"
 6726. 7073
 /gene="VP0309"
 6726. 7073
 /note="similar to GB:AAF95687.1 (AE004322) percent
 identity 64 in 115 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="BAC58572.1"
 /db_xref="GI:28805292"
 /translation="MOHLIFVYGLTARKGEVNNHLLSQAQFLGHEHDAQFSLYDVGSY
 PLASGQRALQGEVYLLDDDTLVLDKIEDVPEVRRRSIATPFGQAMIVLYQDTEQL
 TREIASGDMQGV"
 7170. 7532
 /gene="VP0310"
 7170. 7532
 /note="unknown"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="BAC58573.1"
 /db_xref="GI:28805293"
 /translation="WRVLIIGAMMLLSACAQTLTPASTVTDQWAGKOSALDGLRE
 LEQERIALKDGTHATPELLIAYQNGYQEGQYCEQSAVYLGVVGPYFGICDDVDP
 FQHDYDAGRMSAGAP"
 complement (7607. 8137)
 /gene="VP0311"
 complement (7607. 8137)
 /note="similar to GB:AAF95686.1 (AE004323) percent
 identity 90 in 176 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1

		/product="inorganic pyrophosphatase"	
		/protein_id="BAC58574.1"	
		/db_xref="GI:28805294"	
		/translation="MSLNHYPAKSLPEDIYVIEIPANADPKIEVDKSAVEVDR	
		EKSAHPFYPCNTGYNNITLSLSDPVDVLPYPLMPSVIRCRPVGLKMTDSGE	
		DAKVAVPSKISKEYEHIQVDVIDPELLKAQITHFPERYKELESKMVKVDMADVE	
		AAKAEIIQSYERQONK"	
gene		complement (8265..3281)	
CDS		/gene="VP0313"	
		complement (8265..3281)	
		/gene="VP0313"	
		/note="similar to GB:AAF95685.1 (AE004323) percent	
		identity 88 in 336 aa"	
		/codon_start=1	
		/evidence=not experimental	
		/transl_table=11	
		/product="fructose-1,6-bisphosphatase"	
		/protein_id="BAC58575.1"	
		/db_xref="GI:28805295"	
		/translation="MSGMRITGEFIVEKQADFPHASGDLSLSIASIRLAKIVREIN	
		AAGSDITGAVGTENYQGAQOKLDVYANDKEKALFADDOYCGVASEDEDAVANK	
		ELNONAKYVYVMDPLDSSNIDVNVSGTIFSIYRVSHPMTIPATEEDPLOGHVOA	
		AGVIVGSSITMLVYTTNGVNGPTVDPISGFRCLSHENMTPEDEKTIYSINGNITRF	
		POGVKVIKYCCQENVPEDEPRTSRTISGLVADPHRLNLKGITVLPSTQSHPOKLR	
		LYECNPMAFLIEQMGIIASDGVNRLMDIKFTELHORVPFVGSNNVAKREFFLEIH	
		RDE"	
gene		9265..9360	
CDS		/gene="VP0312"	
		9265..9360	
Query Match		36.2%; Score 1052.4; DB 1; Length 298900;	
Best Local Similarity		76.6%; Pred. No. 0;	
Matches 1303; Conservative		0; Mismatches 391; Indels 6; Gaps 1;	
QY	2	AGNAGGTTTGTATAGCTCAGGCTATATGCAAGTTTCCCAATATTGACCCGCTATTGTTT	61
DB	240005	AGAGATGATTTATATGCTCAGGGAATACCTCAATTTCCCAATATTGATCCGTTTATGTC	240064
QY	62	TGATGAGCCCTCTAGCGGTGCGTGTGTATGCTTGATGATTTGATGGTTTCTTTT	121
DB	240065	TCTATTGCTCCTGTTCAATGCTGTGTACGCTTTGATGTACCTGTCGTTTCATGTTT	240124
QY	122	GCTATGTGTGCAATGCGGAGCGATGCGGGGCGAGGTGTGACCGGTAGCA	181
DB	240125	GCATGTGTGTGCGCAACCTGTGTCTGCAAACTTGGAGCGCTGTGATCTGAGAGCA	240184
QY	182	GTCTGTACTGTATTTGCGCGGCTTTTATGAGTGTATGATGCGTGTGCGAGTTGTTAT	241
DB	240185	GTCTGTACTGTATTTGCGCGGCTTTTATGAGTGTATGATGCGTGTGCGAGTTGTTAT	240244
QY	242	GTGATCTTCAATTTTGTATCTGTCTGTGACCTCTTTATTATTCAAGTGTG	301
DB	240245	GTATCTTCTATTAATTTGAACTGTTTTAGATGATCTTTGTATCTTCAAAATTTGG	240304
QY	302	ACTGCGGATGTCTTCCACCGCGCTTATTTGAGTGTATGATCAAGCCATGTTCTGTAT	361
DB	240305	ACTGCGGATGTCTTCCACCGCGCTTATTTGAGTGTATGATCAAGCCATGTTCTGTAT	240364
QY	362	GCGCGTAAACCAACGACCTTTCTTTGTGTGCGCAATTTTGTGCCCTTTAGTGCA	421
DB	240365	GCAACAAGAAAGTGTGAATCTTTCTTTGTGTGAGCACTTTGTGCCCTTTAGTGCA	240424
QY	422	TTGCGTTTGGGATGGAGATCGTACCTTATGATGATGAACTTTGGGAGAGATA	481
DB	240425	TTTGGCTTGGGATGGGCGGATGTGATGATGATGAACTTTGGGAGAGATA	240484
QY	482	ACGATGTGCTTGGGCTTTTGTATTTCCCTAATGTGCGCACTGCGCGCATCTTCA	541
DB	240485	ACAGATGTGCTTGGGCTTTTGTATTTCCCAATGTGCGGCTTTTACCACTGCACTG	240544
QY	542	CAGCTTATGAATCGCTTGAAGCGGTGTTCTTCTTATTTCTTAATTTGTTTAT	601
DB	240545	CAGCTTATGAATCGCTTGAAGCGGTGTTCTTCTTATTTCTTAATTTGTTTAT	240604
QY	602	GCTAAACCTCGTCCGCTAGCGAGGTATCCGACTGTTTTAGCTGATGATGATTC	661
DB	240605	AAAAAGCCTCGCCCAATGCGTGTGCGTATCAGGGCTATTTTATGCTGATATGATTC	240664
QY	662	CGCTTCTTGTGATAGTCCGTGAGCCAGATGCTGAGTGGGTCTGTTGGCGCTTC	721
DB	240665	CGCTTCTTGTGATAGTCCGTGAGCCAGATGCTGAGTGGGTCTGTTGGGTGTTAT	240724
QY	722	ATTTCATGGGCAATCTCTCTTACTATGATGATCATCGATTTTGTATGATGTT	781
DB	240725	ATTTCATGGGCAATCTCTCTTACTATGATGATCATCGATTTTGTATGATGTT	240784
QY	782	TGCTTTCAAGCCCGTTTGTATCAAGAC-----CGTGAAGCAAAATAGGTTAGT	835
DB	240785	TGGCTTCAAGCCCGTTTGTATCAAGAC-----CGTGAAGCAAAATAGGTTAGT	240844
QY	836	TAGGTAAACAGTATTTAGATCTTTGTACGCGCATCGATCAAGGTTTGGGTTGAA	895
DB	240845	GGTGAAACAGTATTTAGATTTTGTGTCAAGCATCGTATCAAGGTTTGGGTTGAA	240904
QY	896	AATGAACGAACGGCAACCGTTGTTTGACTGTATTAATGCCATTTGACCTACATGTG	955
DB	240905	AATGAACGAACGGTAAAGCGTTGCTGACCGTATCAACCGGATTTAACTACATGTT	240964
QY	956	GGCAACATCAGTTTCTCTAGTACTACACGGAAGTTTGGAAAGCTGCCGTAGCC	1015
DB	240965	GGCAACATCAGTTTCTCTAGTACTACACGGAAGTTTGGAAAGCTGCCGTAGCC	241024
QY	1016	GAGTTGCTCGCTATTTGTTGTGTTACGATTAATGCGCGGATTTTCCCAATTAGTACC	1075
DB	241025	GAGTTGCTCGCTATTTGTTGTGTTACGATTAATGCGCGGATTTTCCCAATTAGTACC	241084
QY	1076	AAAACCTGGGATGCTAATGCAATTTAAACCAAGCATGCTCAACAATCTTACCGTAA	1135
DB	241085	AAAACCTGGGATGCTAATGCAATTTAAACCAAGCATGCTTGTGATTAACCATACCGTAA	241144
QY	1136	GGTAGAGATGATGAGGAGCGGTGTATGTTTCAAGGTTAGACTTGGGCTTAAAGCTGAT	1195
DB	241145	GGTAGAGATGATGAGGAGCGGTGTATGTTTCAAGGTTAGACTTGGGCTTAAAGCTGAT	241204
QY	1196	GGTGTATATTTGACCGATTGAAAAAGATTGTTGATGATTTGACCGTGTGATGAC	1255
DB	241205	GGTGTATATTTGACCGATTGAAAAAGATTGTTGATGATTTGACCGTGTGATGAC	241264
QY	1256	CGAGGTGAATTTCTTAATCTTCAATCCGGGTGATTTCAATGGGGTGTGGCGCCT	1315
DB	241265	CGTGTGAATTTCTTAATCTTCAATCCGGGTGATTTCAATGGGGTGTGGCGCCT	241324
QY	1316	TGATGTACAGCATATTTTCACTAGTGGGAGTACCTTGTATCTCAACAGTACTAG	1375
DB	241325	TGATGTACAGCATATTTTCACTAGTGGGAGTACCTTGTATCTCAACAGTACTAG	241384
QY	1376	CGTTATGATGATGCGCTTGGGTTGAATTTAAATGTTGAGTTTATGTTTCTT	1435
DB	241385	CGTTATGATGATGCGCTTGGGTTGAATTTAAATGTTGAGTTTATGTTTCTT	241444
QY	1436	GCGCTGATGACAGATCAAGGAAAAAGCCGGGCTTGGCTGATACAAAGTCTCAAT	1495
DB	241445	GCGATATGCGGCAATTTACGGGCAAGAGCCAGGCAAGCATTCATTAATCTGTTAAC	241504
QY	1496	GCGCACTTTTACCAATCACTGCAATTTGATGCGGATGTGAGCTTAAACGTGACCA	1555
DB	241505	GCGCACTTTTATGAATCACTGCTTGAAGTGAAGCGGATTTGAAGCGTGAACCG	241564
QY	1556	TTCCAGCGCTCAGTTCCATATCAATCCAAAGATTTAAACTGACAGATTTGAAACT	1615
DB	241565	TTTCAAGCCCAACATTTCAATTCATTCGAGATTTAAATCTTGAAGATTTGAAACT	241624
QY	1616	TGGGTCACTTTGATGATTTTGAAGTCAACCGATATCAAGTTCCAGATCTTATCAATAC	1675
DB	241625	TGGGTCACTTTGATGATTTTGGGATGAAGGATGAAGACCAACAGATCAATCTTAT	241684

/db_xref="GI:27359909"
/translation="MSGFENLIEKYFVGROAKRQXVHLAAGDCAIYKPSNVQIAIS
TPTLVCGTFLADADPAWVAHKAASINISDLAMGATPMAVSFALMLPADAMAP
CGGFLADYFGILIGDGTGKPLSTMTTNGCPBEKALITRSAAKEDYFPTGT
GAKKALDVIILSNLRTPRPAQALEKAMEMPRILAGALVLSAIDISGLAD
LGHILRSQVSVSLDYSKLPISSELSFLITEREMALQVALTSGEEYELCTVPEBARG
TLESALAHVCVCYTCIGQIRPHGFELHDDGTKLWDLSDGFHFA"
/complement(6847..7314)
/locus_tag="VV10318"
/complement(6847..7314)
/locus_tag="VV10318"
/note="COG0781"
/codon_start=1
/transl_table=11
/product="transcription termination factor"
/protein_id="AAO08848.1"
/db_xref="GI:27359910"
/translation="MGASYKPAARPAARPAALQAIYSNQTENATIEAOFAGEKY
DEBRRAAPALAPETDVAFRDILTVLSHTLSDSKIRYVSFPQDLMELAL
LRLAVEMTRREDVPEKVINIEALIKAFVAFEDSHKFNVGLDKAAHVERK"
/complement(7314..7784)
/locus_tag="VV10319"
/complement(7314..7784)
/locus_tag="VV10319"
/note="COG0054"
/codon_start=1
/transl_table=11
/product="Riboflavin synthase beta-chain"
/protein_id="AAO08849.1"
/db_xref="GI:27359911"
/translation="MKVIEGEPAPNAKIAIVISPNFINSILSGAIDTLKHGOY
SEDNITVPCGAVELPIYACQVAKTKRDAIVSGSTRGTHFEDYCEMNKGLA
QVLSFSLPVARCVLFDITDQATERRAGKGNKAALSALEMINVLSIDS"
/complement(7961..9070)
/locus_tag="VV10321"
/complement(7961..9070)
/locus_tag="VV10321"
/note="3,4-dihydroxy-2-butanone 4-phosphate synthase;
COG0108; COG0807"
/codon_start=1
/transl_table=11
/product="GTP cyclohydrolase II"
/protein_id="AAO08850.1"
/db_xref="GI:27359912"
/translation="WPISTBOHIEDIRLQKVMILMDEBENEDILMAEHITPERA
INEMKXRGULCTMTKERCRRLCPMVDNNNAQVITNTFVSI EAEVYTTISAA
DRAITVQAAVAKAFAADLVOPGHLPLAADGVLTAEAGTGAECCLARLAGEPBS
VARELNDGTAARRPDLVEAEQHLGLQIDILIEVNNTEETIERVACXKPTFE
GEFDLVTRDITDQIHFLARKGTGDDTPVRLDPTFDLRSDBNARSWTDLK
AMKRIQGGGVVILIGNESEPELLIHRVMEADQKEAPRLAKKQGSRRVVGSDI
LADLVHMDRLISSNSKKYHALGFGNLVVEYCE"
/complement(9168..9824)
/locus_tag="VV10322"
/complement(9168..9824)
/locus_tag="VV10322"
/note="COG0307"
/codon_start=1
/transl_table=11
/product="Riboflavin synthase alpha chain"
/protein_id="AAO08851.1"
/db_xref="GI:27359913"

Query Match 35.6%; Score 1035; DB 1; Length 301442;
Best Local Similarity 75.9%; Pred. No. 3e-309;
Matches 1278; Conservative 0; Mismatches 405; Indels 0; Gaps 0;

QY 8 TTTCCTAAGGCTTACGAGTTATCGAGTTTCCCATATTTGACCCGCTATGTTTGCATC 67
DB 215617 TTCATTAGTCTCAAGGTTACTGCTTCTTCTATATGACCCCGCTCTTTTCTCGATA 215558

QY 68 GGCCTCTAGAGCGTGCCTGTGATGCTGATTTGTGGGTTCTTTTGTGCTATG 127
DB 215557 GGAACCATCTAGTCTGTGGTACGGGTTGATGATCTATTTGGTTTCTTTTGGCAATG 215498

QY 128 TGCTTGGCCAAATCCGCCAGCGGATCCGCCGGGAGCTGTTGACCTGCTAGCAATCTCT 187
DB 215497 TGGCTCGCAATTCGCCAGCAAGAAACCGGGAGTGGCTGGAGCCAGAACAAAGTTTGC 215438

QY 188 GACTTGTATTCGCGCGCTTTTATGATGATGATCGGTGGCCAGTGTGTTATGATC 247
DB 215437 GACTTGTGCTTCCGCTGATTCCTAGGGCTTGTGCTGGTGTGCTCCATTTGGTACGTCTG 215378

QY 248 TTTCACAAATTTATCTGTTCTTCTGCTGACCCCTTTTATTTATCAAGTGTGACTGGC 307
DB 215377 TTCTATAACTTCGATCTCTCTCTGCTGATCCGATTTTACCTCTTAAAGTCTGACAGGT 215318

QY 308 GGCATGTCCTTCCAGCGGCGCTTATGAGGTGATCAACCGCAATGTTCTGGTATGGCGGT 367
DB 215317 GGCATGTCCTTCCAGCGGCGCTTATGAGGTGATCAACCGCAATGTTCTGGTATGGCGGT 215258

QY 368 AAAAACCAGACGACCTTCTTGTGTGGCCGATTTGTGCCCCCTTTAGTGCATTCGGT 427
DB 215257 AAAAATGGCGAACCTTTTGTGGCGTGGCGGACTTCGTTAGTGCATTTGGC 215198

QY 428 TTGGGATGGGAGATCGATCGATCTTATGATGATGATGATGATGATGATGATGATGATGAT 487
DB 215197 TTAGGAGTGGGCGCTTAGGCACTTATGACCGGAGCTTTGGGCGCAAGTACCGAT 215138

QY 488 GTGCTTGGGCTTTTGTATTCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 547
DB 215137 GTGCTTGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 215078

QY 548 TATTAATTCCTTGAAGAGCGGTGCTGCTTTTATTTTATTTATTTATTTATTTATTTATTT 607
DB 215077 TATGAATGACATTAAGAGGAGGTGCTGCTTTTATTTATTTATTTATTTATTTATTTATTT 215018

QY 608 CCGTGTCCGCTAGGACGATCGGACCTGTTTATGATGATGATGATGATGATGATGATGATGATGAT 667
DB 215017 CCAAGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214958

QY 668 CTGTGGAATACGTCGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
DB 214957 CTGCTGAATACGTCGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214898

QY 728 ATGGGGGAATCCCTCCCTTACCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
DB 214897 ATGGGACAGATCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214838

QY 788 TACAAGCGCGTTGATCAAGACCGGTGACGACAAATAGAGGTAGTTAGTGAACAG 847
DB 214837 TATAAAGTGGCAGTATCAAGACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 214778

QY 848 TATTAGATCTTGTCAAGGCAATCGTCGATCAAGGTGTTGGGTGAAATGACGACAG 907
DB 214777 TATTGAATATGCGCGTGCATCTGTGATGAAGGCGACCTGGGTAAATGACGATAC 214718

QY 908 GGCAGAGGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
DB 214717 GGTAAAGGCGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214658

QY 968 TTTCCTCTAAGCTACAGGCAAGAGTTTGGAAAGCTGCGGTACCGAGTTGCTGGC 1027
DB 214657 TTTCCTTGTGTTACAGCGGTAAAGAGTTTCTGAAAGCGCGGTGCAAGAGTTGCTGGC 214598

QY 1028 TATATTCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
DB 214597 TATATTCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214538

QY 1088 GCTAATGCCAATTTAAACCAAGATGCTCAACATCTTACCGTAAAGGTGAGATGAC 1147
DB 214537 GCAAGCGCAATCTCAATGATGCTTGTGATTAACCAATCTGCAAGGCAAGATGAC 214478

QY 1148 ATGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
DB 214477 ATGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214418

QY 1208 GACCACTGAAAAAGATTGTTGATGATTTGAGCCGTGCGCTTATGACCGAGGTGAATT 1267
 Db 214417 GACCAATTTGCGTAAATTTGTCGATGACCTGACCGCGGAGTACATGACCGGGGAGATC 214358
 QY 1268 CTTAATCTTCAAAATCCGGGATGATTTCAATGAGGGGTTGGCCCTTGATGATGACG 1327
 Db 214357 CTCATTTCTTCAAAATCCGGGATGATTTCAATGAGGGGTTGGCCCTTGATGATGACG 214298
 QY 1328 CATCATTTTTCATTTGCTGGGGGATACCTTGTATCTCAACATGACTAGCGTTGATGAT 1387
 Db 214297 CATCATTTTTCATTTGCTGGGGGATACCTTGTATCTCAACATGACTAGCGTTGATGAT 214238
 QY 1388 GTGCCCTTGGGGTGAATTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 1447
 Db 214237 GTGCCACTGGGGTGAATTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 214178
 QY 1448 CAGATCAGAGGAAAAAGCCGGGCTTGGCTATGACAAAGATGCTCATGCGCATTTTC 1507
 Db 214177 CAGATCAGAGGAAAAAGCCGGGCTTGGCTATGACAAAGATGCTCATGCGCATTTTC 214118
 QY 1508 CAGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1567
 Db 214117 GAGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214058
 QY 1568 CAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1627
 Db 214057 ACGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213998
 QY 1628 GATGATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687
 Db 213997 GATGATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213938
 QY 1688 TAA 1690
 Db 213937 TAA 213935

RESULT 10
 AP005332 249150 bp DNA linear ECT 04-DEC-2003
 LOCUS Vibrio vulnificus YJ016 DNA, chromosome I, complete genome, section 3/14.
 ACCESSION AP005332 BA000037
 VERSION AP005332.1 GI:37197435
 KEYWORDS
 SOURCE
 ORGANISM
 Vibrio vulnificus YJ016
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.

REFERENCE 1
 AUTHORS Chen, C.Y., Wu, K.M., Chang, Y.C., Chang, C.H., Tsai, H.C., Liao, T.J., Liu, Y.M., Chen, H.J., Shen, A.B., Li, J.C., Su, T.L., Shao, C.P., Lee, C.T., Hor, L.I. and Tsai, S.F.
 TITLE Comparative Genome Analysis of *Vibrio vulnificus*, a Marine Pathogen
 JOURNAL Genome Res. 13, 2577-2587 (2003)
 REFERENCE 2 (bases 1 to 249150)
 AUTHORS Chen, C.Y., Wu, K.M. and Tsai, S.F.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-2002) Shih Feng Tsai, National Health Research Institutes, Division of Molecular and Genomic Medicine; 128, Yen-Chiu-Yuan Road, Sec 2, Taipei, Taiwan 115, Republic of China (E-mail:petesai@nhi.org.tw, Tel:866-2-8146-1041, Fax:866-3-2789-0484)

COMMENT This sequence was determined by the Sequencing Core of the National Yang-Ming University Genome Research Center (YMGRC),
 http://genome.ym.edu.tw/
 Location/Qualifiers

FEATURES
 source
 1..249150
 /organism="Vibrio vulnificus YJ016"
 /mol_type="Genomic DNA"
 /strain="YJ016"
 /db_xref="taxon:196600"
 /chromosome="I"

gene 299..979
 /gene="VV0511"
 CDS 299..979
 /gene="VV0511"
 /note="identified by Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93275.1"
 /db_xref="GI:37197436"
 /translation="MNVSRKKNKRLTTRKTFGLRPIQTEHGPLYLDYLRKMTNID
 IADYRPLMAIRVDLRPEKLRKMSGNVMTDFLRISQSIDHSGKKRGGVVRH
 PCVRLAVRERSSSINDYHVLVFNDRFNVWLGKTQOENLSSFLVEAMAVAGV
 DYRANGLVHPSKRXNSDSVTHRLNCSDAFDDAFDFPFKHISYLAKENTKRGSGR
 RNRGSHK"
 865..888
 /gene="rpt112"
 865..888
 /gene="rpt112"
 /note="2 units of 12 bp in-tandem repeat"
 /rpt_type=tandem
 /rpt_unit="TGATGCGTTGA"
 1040..1708
 /gene="VV0512"
 1040..1708
 /gene="VV0512"
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93276.1"
 /db_xref="GI:37197437"
 /translation="MTHQINVAHQRAHTTHDRFNGKPYFYFEDGLVLEYLEIDY
 LTELNKYRITVAHINLPSDFGDDYLVPAHFRILLESFBLCKDRTDYTYR
 QQTVIRYITWKASBSSTSDHHLILFDRDLPGTGEGKGRFLYKRVKAMSVVE
 AYYRQPSGLGYSPNREGFELSDAEVFPLOINFEFYRISRLAKPITNLDRKXSG
 CSYHD"

gene 1286..1441
 /gene="VV0513"
 CDS complement(1286..1441)
 /gene="VV0513"
 /note="identified by Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC93277.1"
 /db_xref="GI:37197438"

gene 1733..2368
 /gene="VV0514"
 CDS 1733..2368
 /gene="VV0514"
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93278.1"
 /db_xref="GI:37197439"

gene 2436..2642
 /gene="VV0515"
 CDS 2436..2642
 /gene="VV0515"
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="predicted transcriptional regulator"
 /protein_id="BAC93279.1"
 /db_xref="GI:37197440"

[illegible][illegible]

QY 1320 TGTACGACCATATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTACGCGTT 1379
DB 482 TGTACGACCATATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTACGCGTT 541
QY 1380 CATGTGATGTGCGCTTGGGGGTGAATTTCAACATGATGAGGTTTATGTTCCCTTGGGC 1439
DB 542 CATGTGATGTGCGCTTGGGGGTGAATTTCAACATGATGAGGTTTATGTTCCCTTGGGC 601
QY 1440 TGTATGGCAGATATCAAGGAGGAGAAAGCCGGGCTTGGCGATACAAAGTCCGATGCGC 1499
DB 602 TGTATGGCAGATATCAAGGAGGAGAAAGCCGGGCTTGGCGATACAAAGTCCGATGCGC 661
QY 1500 ACATTACCAAGATCAATCGAATTGATGCGCGATGTGAGCTAAAGGTGAGCCATTC 1559
DB 662 ACATTACCAAGATCAATCGAATTGATGCGCGATGTGAGCTAAAGGTGAGCCATTC 721
QY 1560 CAGCGCTCAGTTCATATCAATCAAAAGATTAACACTGCAAGATTGGAACTTGGG 1619
DB 722 CAGCGCTCAGTTCATATCAATCAAAAGATTAACACTGCAAGATTGGAACTTGGG 781
QY 1620 TCACTTGGATTTTGAAGTCAAGGATATCAAGGATCAAGTCAAGTCAAGTCAAGTCAAGT 1679
DB 782 TCACTTGGATTTTGAAGTCAAGGATATCAAGGATCAAGTCAAGTCAAGTCAAGTCAAGT 841
QY 1680 TTTCAAGTCTAA 1690
DB 842 TTTCAAGTCTAA 852

RESULT 12
AX009482 838 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 2 from Patent WO961634.
DEFINITION AX009482
ACCESSION AX009482
VERSION AX009482.1 GI:9996767
KEYWORDS
SOURCE Vibrio cholerae
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE 1
AUTHORS Carlin, N. and Lebens, M.R.
TITLE Method of producing thy a<->strains of vibrio cholerae, such
JOURNAL strains and their use
CARLIN NILS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
FEATURES
source 1. 838
/organism="Vibrio cholerae"
/mol_type="unassigned DNA"
/db_xref="taxon:666"

ORIGIN
Query Match 28.8%; Score 838; DB 6; Length 838;
Best Local Similarity 100.0%; Pred. No. 1.9e-248;
Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TGTATCTTCTACATTTTGTATCTGTTCTTGTGCTGACCCCTTTTATTATTCAAGTGTG 300
DB 241 TGTATCTTCTACATTTTGTATCTGTTCTTGTGCTGACCCCTTTTATTATTCAAGTGTG 300
QY 301 GACTGCGGAGTGTCTTCCACGCGGCTTATGGGATGATACCGGATGTTCTGTGA 360
DB 301 GACTGCGGAGTGTCTTCCACGCGGCTTATGGGATGATACCGGATGTTCTGTGA 360
QY 361 TGGCGTAAACCAACGACCTTCTTGTGATGCGGATTTTGTGCTTGTAGTGC 420
DB 361 TGGCGTAAACCAACGACCTTCTTGTGATGCGGATTTTGTGCTTGTAGTGC 420
QY 421 ATTGCGTTGGGAGTGGAGTATCGGTATCTTATGAAATGAACTTTGGGAGAGT 480
DB 421 ATTGCGTTGGGAGTGGAGTATCGGTATCTTATGAAATGAACTTTGGGAGAGT 480
QY 481 AACGATGTCCTTGGGCTTTTGTATTCCTTAATGATGCGGACGCGGATCTTC 540
DB 481 AACGATGTCCTTGGGCTTTTGTATTCCTTAATGATGCGGACGCGGATCTTC 540
QY 541 AACGATGTCCTTGGGCTTTTGTATTCCTTAATGATGCGGACGCGGATCTTC 600
DB 541 AACGATGTCCTTGGGCTTTTGTATTCCTTAATGATGCGGACGCGGATCTTC 600
QY 601 TGTAAACCTCGTCCGCTAGGAGCGATATCGGACTGTTTGTAGTGAATACGGTACAT 660
DB 601 TGTAAACCTCGTCCGCTAGGAGCGATATCGGACTGTTTGTAGTGAATACGGTACAT 660
QY 661 CCGCTTCTTGTGAATAGTCCGTAGCGGATGCGGATGCTCAGTTGGGTGTGGTGGCTT 720
DB 661 CCGCTTCTTGTGAATAGTCCGTAGCGGATGCGGATGCTCAGTTGGGTGTGGTGGCTT 720
QY 721 CATTTAATGGGCAATCTCTCTTACTATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 CATTTAATGGGCAATCTCTCTTACTATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 TGTGCTTACAGCGGCTTGTATCAAGACCGTGTAGCAGCAAAATAGGATGATG 838
DB 781 TGTGCTTACAGCGGCTTGTATCAAGACCGTGTAGCAGCAAAATAGGATGATG 838

RESULT 13
AE006043 10719 bp DNA linear BCT 08-MAR-2001
LOCUS AE006043/c
DEFINITION Pasteurella multocida PM70 section 10 of 204 of the complete
genome.
ACCESSION AE006043
VERSION AE006043.1 GI:12720282
KEYWORDS
SOURCE Pasteurella multocida
ORGANISM Pasteurella multocida
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Pasteurella.
REFERENCE 1 (bases 1 to 10719)
AUTHORS May, B.U., Zhang, Q., Li, L.L., Paustian, M.L., Whitlam, T.S. and
Kapur, V.
TITLE Complete genomic sequence of Pasteurella multocida, PM70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
PUBMED 11248100
REFERENCE 2 (bases 1 to 10719)
AUTHORS Zhang, Q. and Kapur, V.
TITLE Direct Submission
JOURNAL University of Minnesota (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA
FEATURES
source 1. 10719
/organism="Pasteurella multocida"
/mol_type="genomic DNA"
/strain="PM70"
/db_xref="taxon:747"

gene
225..1076
/gene="PM0074"
CDS
225..1076
/gene="PM0074"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02158.1"
/db_xref="GI:12720283"
/translation="MKSEESTCTCVSPPTKMAQIGEDVGVYKATKQIYSPFAIPAG
AFIAIPFVATQONGDAPMGLTKLVGIVFSLGIYVVCSELEFSTSTWTVAR
VSGRIPTROMLRNLYVYCGNFVGLFVLTTFAGQIMANQOMGLTIKTQKH
HTWDEALGICINIMVCIATWMTYAKTILLDKAVIMILPISMFVAGSPHSAVMY
IPGMMLIHFAFSPFPMQVSLNPAQFADLMTHTFVYNLLPVTIGNVGNLCVALTO
WYINRPH"
gene
1156..3486
/gene="pflB"
/note="synonym: PM0075"
CDS
1156..3486
/gene="pflB"
/codon_start=1
/transl_table=11
/product="pflB"
/protein_id="AAK02159.1"
/db_xref="GI:12720284"
/translation="MSMSQLMETQCKAMEFTGGDMOTENVVRDPIQKNYTPYEGDES
FLADTERITKLMNDWEKIKYENKTKHPYDIDDDTSTISHAPGYIDSLKIEVL
QTDALKRAIMPFGEINWKSCKVTRRELPEYBQIFTEYRKHNGVDPVYTDIL
RCRKGVIITGLPDAGRGRIIGDYRPMALYGADFLMKDKNQVOTSLQDKERGDIDA
TIOLEEEIAEOHRAFGKMEWASGYDISGPATNAEAVQWTFAYLAIVKSONGA
MSFGVSTFLDIYIERDLKAGITEQEAQELIDHLWMLRMVRELRTPEYDQFLSGB
WMAETLAGMGLDRLTWNKSRFRIHTLYMGSPSEPNLITLSEKLPBEGKYAAK
VSDTSSVOYENDLMRPPFONDVYAIACVSPMVGKQMPFGARMLAKTLLIYAN
GVDSKSDQVQPKTDPTITSELDVDPMTLSDMLAKOYATALNITHFMKDVA
YEALMALHDRVFRITKAGIAGLSVADSLSAIKYAKYVRGDIETIKKAGVQV
AKDVAIDIEIEGEYPOFGNNDNRVDEIACDIVERPKTIQKLGITRNATPTOSLTIT
SNVVGKTKTNGPDGRSGPFGANPMHGRDKGAVASLTVAKLPFAIAKQGISY
TSIVNALGKDYEAQKRNLAGMDGYFHHEATIEGGQHLNWNVNRNEMLDAMENE
KYPOLITRVSYAVRNSLTKEQQDQVITRTFTQAM"
gene
3659..5698
/gene="est"
/note="synonym: PM0076"
CDS
3659..5698
/gene="est"
/codon_start=1
/transl_table=11
/product="Bst"
/protein_id="AAK02160.1"
/db_xref="GI:12720285"
/translation="MKINTLSFALFCSLSAALAKADVVPFGDSJSDMGQCTGNKKA
LKADGSYHLIDYELAQAFGKLIPTQGLNVAYSGVIVGANTRTAEQPHIALEK
QINEYLAHVKKKALHILWAGNDLATALVATVKTPEEKOAYVLASINDMAQTMQ
OMGALQOAGVNOIIAPTIENVTYPEPFDKGEAAGIOAKSYGLIKOSDFVANFKA
AAQOLUOETRLREPERKHROTLEKAEDPFNSRYFTQIALSAGYNAGIAETLI
KEKTIVEQAOATAFNAISITSALNQVQGNIVRTDIDGILKDMMTTAEYGLNNTT
LVCKESTAPAPACKPEPDQNASORLEQOQSAQSVDTLAAVQKQSHSLHAKQVQFN
LQIVQOQTEALADPTRESNHRLOQOQSAQSVDTLAAVQKQSHSLHAKQVQFN
OMQALVAVQOQKQVQGLININTONRYVISTALRYDADNMWLGVLQVNNNTLTDRY
AYIGHAHOQSAETELESFSGMLPAGYEKFKALGIALIADIKTKTQVNAFGENRG
ITQMCBAARTKESVSGVGFDRYQASWQPVYTRVYEMNNDIPTIOALINSSEFM
TUTPSGDNONVNI VAGLHKPISSGFHNLAIRDLGRNMLSNITTLQAGIGTFP"
gene
5862..6602
/gene="act"
/note="synonym: PM0077"
CDS
5862..6602
/gene="act"
/codon_start=1
/transl_table=11
/product="Act"
/protein_id="AAK02161.1"
/db_xref="GI:12720286"
/translation="MAVVGRIHSYSCGTVPDGRIRFILFMQGLMRCQYCHNRDITW
LHAGKEITVEELMKKEVVVYTRHFMNASGGGVATSGEAVLQAEFVADMWSACKKEGHT

gene
CLDITNGFVRNVDHVIDELDDVTDVLIDLKQNDKRIHQNLIGVNRKTLERAOYLAKR
NOPWIRYVYVPGYTDADDEDIHILGHFIKDKKNIEKVELLPHRLGAKHKBAMGEKYE
LEEVNPTKSELEHKKSLSESGHIIK"
gene
complement (6819..7319)
/gene="PM0078"
complement (6819..7319)
/gene="PM0078"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02162.1"
/db_xref="GI:12720287"
/translation="MQSDMDKEMRHLRLADKAEQGEIVGAVLVDEQGNLGEW
NLISIEDPTAAIEIVALROGQORLONRLINTLYTLBPCQWCAQILSRIRLV
PQADYKTVAGVSRPHFPDDYKMHNAIRITAGVLOECSQKISAFQKRRBOQKIAKO
KSHLFT"
gene
complement (7334..8185)
/gene="thyA"
/note="synonym: PM0079"
complement (7334..8185)
/gene="thyA"
/codon_start=1
/transl_table=11
/product="thyA"
/protein_id="AAK02163.1"
/db_xref="GI:12720288"
/translation="MKHYLELCQRIVDEGVVENARKRCCTVIDADLTVDVANNQ
PLITTRKSYWKAIAIEFLGYIRGNVADFRGLGTPTVDANANENSAALNPHRGTD
DMGRVYGVQGRAMRKXENGETVDQIKRIVNNLNGIDDBGEINTFNPGEFELGCLRP
MHTFVPSLIGDLTYLISYORSQDVLGNFNQIOVFETLAMAQITGKAGAKYKRII
NNAHIEDQDLKKNQVLRDPPLELQIHINPEIKTLEDELETWVTMDPKVTGYQCHDA
IKRPSV"
gene
complement (8182..8994)
/gene="lgt"
/note="synonym: PM0080"
complement (8182..8994)
/gene="lgt"
/codon_start=1
/transl_table=11
/product="lgt"
/protein_id="AAK02164.1"
/db_xref="GI:12720289"
/translation="MQSNYFSPFQDPVFEFIEGPIGLRWYGLMILGFLPRLMAVR
ANOTSGMNTDQVDSLFPNGEMVGLRGIGVLPYQFQYPIQDPAVYLFVWEGGMSF
HGGILGVIITAMLITAKTORQGWOTADPAPLIPGLGMRIGNFINDELWGTVDVP
WAVLPSPGSLPFRHSQLYEAVDEGIVLPILNWTIKKRPICATAGLILGYGIFRF
YVFFRPDQGLYFGQHSWGLSTPMILIGAVIMLVAYQSAKREIL"
gene
complement (9004..9798)
/gene="PM0081"
complement (9004..9798)
/gene="PM0081"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02165.1"
/db_xref="GI:12720290"
/translation="MFTFLGICAVGAFAGLAGLFGIGGLVIVPVLYVLMKRYGV
PEMAWACALGTSFATIVTFLFISCSAQRHKLGNVWDVYKAVPMITTFISGFEIS
YDKDLSARI PALIVYILAIKMTLSKPEVTKALITKASTIAGLIGASSAAGIGG
GGFIVPFLNSRGIDMKKATGSSAVQSSLLAIKMSFPIAKKRVGLPEYSNGYVLP
ALVIGITLTSIFTSIGATVTSILPVLKKAFLMLILAINMLK"
gene
complement (9801..10397)
/gene="PM0082"
complement (9801..10397)
/gene="PM0082"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02166.1"
/db_xref="GI:12720291"
/translation="MTDPDGYRPNVGIIVICNSKQVLMARKRYGQNSWQFPOGGINDNE
SAGQARYELFERVGLSPKDVKILYISKMLRYKLPKRLRYDSKVCIGIGQGRMFL
QVAVDERKNIQSSKSPEDGRWVFWIPYQVYSFKKEVYKAKKEFASVLPDQAK

ORIGIN

ENLSSKSNESDLKTHHTYKXSTPLTKHSHKHFHRSK"

Query Match 28.0%; Score 814.8; DB 1; Length 10719;
 Best Local Similarity 68.7%; Pred. No. 5,2e-241;
 Matches 1161; Conservative 0; Mismatches 507; Indels 22; Gaps 2;

4 AAGGTTGGTATGCTCAGGGTTATCTGCAATTTCCCAATATGACCCCGATGTTTC 63
 Db AAGGATTTTATGCAATGCAATTTATTTTCAATTCACAAATTTGATCCCGATTTTGA 8945
 Qy 64 GATGGCCCTTAGCGGTCGCTGATAGCTTGAATGATTTGGGGTTTCCTTTTGC 123
 Db 8944 AATCGGCTTATCGGTTTACGTTGATGCTTAATGATCTGCTGGGGTTCTTTGC 8885
 Qy 124 TATGTTGGTGGCAATCGCCGATCGCGGCGAGTGGTGGACCGTGAAGCAAGT 183
 Db 8884 CCGTTGGCTTGGCGGTAACGCTAATCAACAGGCAATGCTGGAACAAGCAAGT 8825
 Qy 184 CTCTGACTTTTATTCGCGGCTTTTATAGTGAATGATCGGTGGCCGAGTTGTTATGT 243
 Db 8824 GATGAGCCGTTATTTAAACGTTTATAGGCGTTTATAGGTGAGCATTTGTTAGCT 8765
 Qy 244 GATCTTCAATTTTGAATCTGCTGCTGACCTCTTATTTATTTCAAAAGTGGAC 303
 Db 8764 GTTGTTTTATCAATTCGATTTTCTTACAGATCTCGGTAATTTATTCGTTTGGGA 8705
 Qy 304 TGGCGGCAATGCTTCCACGCGGCTTATGAGTGTGATCAACGCCATGTTCTGGTATGC 363
 Db 8704 AGCGGTAATGCTTTCCATGATGCTTAATGATGATTTAGCATTTATCATGATCAGAC 8645
 Qy 364 GCGTAAAAACAAGCACTTCTTGTGTGTGTCCTATTTTGTGCTCTTATAGTCCATT 423
 Db 8644 TAAATATCAAAAACGCGGTTTCTGGCAACGCGGATTTTGGCGCGGTGATTCATT 8585
 Qy 424 CGGTTGGGATGAGGACGATCGGTAATTTATGATGATGATGATTTGGGACAGATAC 483
 Db 8584 TGGCTTAGTATGAGACGATCGGCACTTTATCAATGACAAATTTGGGGCGGTATAC 8525
 Qy 484 GATGTCCTTGGGCTTTTGTATTCCTTAATGATGAGTGGCCCACTGCGCGCATCTTCA 543
 Db 8524 GATGTCCTTGGGCTTTTGTATTCCTTAATGATGATGATTTTACCAACCATTCATCA 8465
 Qy 544 GCTTATGATTTGCGCTTAGAGGGTGTCTGTTCTTATTTCTTAATTTGTTATG 603
 Db 8464 GCTTATGATTTGCGCTTAGAGGGATTTGATGATTTCTTATTTCTGAATTTGATATCA 8405
 Qy 604 TAACTCTGTCGCTAGGCAAGGATTCGGGATGTTTAACTGATGATACGATATTCG 663
 Db 8404 AAAGCTCGACCAATCGGCGCAACCGCTGTTTATTTCTTACGCTGATCGGATTTTCG 8345
 Qy 664 CTTCCTTGTGAAATACGCTGAGGCAAGTCAATGTTGGTCT---GTTTGTGCTT 720
 Db 8344 TTTTACGTTGATTTTTCGAGCTGATGCTCAATTTAGGCTTATTTTGTGCAACA 8285
 Qy 721 CATTTCAATGAGGCAATCTCTCTTACCTTACCTTATGATGATGATGATGATGATGAT 780
 Db 8284 TATCTCAATGAGCAAAATCTCTCCACCAATGATTTCTCATTTGCTGGGATATATCTT 8225
 Qy 781 TTGCTTTTCAAGCGCGGTTTGTATCAAGCCGTAGACCAAAATGCGGTAGTTAGT 840
 Db 8224 AATGCAATATCAAAATG-----CGGCTAAAAAAGGCAAAATATAT 8184
 Qy 841 GAAAGATTTATGATCTTTGTCAGGCAATGTCATCAAGTATTTGGTTGAAATATGA 900
 Db 8183 GAAACATTTATGAACTTTTGCACATGATCTTATGAAAGTGTGTTGGTTGAAAAATGC 8124
 Qy 901 ACGAAAGGCAAGCGTTGTTACTGTGATTAATGATGATGATGATGATGATGATGATGAT 960
 Db 8123 ACGGATGCAAAAGCTGTCTGATCGCTGATGATGATGATGATGATGATGATGATGATGAT 8064
 Qy 961 CAATAGTTCTCTAGTACTACAGCAAGGTTTGGAAAGTGGCGATGACCGGATTT 1020

Db 8063 TAAACAATTTCCGTGATCAACACAGCAAAAGCTATTTGAAAAGCGCAATTCAGATTT 8004
 Qy 1021 GCTGGCTATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 Db 8003 TTTTGGTTATATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7944
 Qy 1081 CTGGGATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 7943 CTGGGATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7884
 Qy 1141 GAT 1200
 Db 7883 GAT 7824
 Qy 1201 TCAATTTGACCACTTGAATAAGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 7823 AACCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7764
 Qy 1261 TGAATTTCTTACTTATCAATCCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Db 7763 TGAATTTCTTACTTATCAATCCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 7704
 Qy 1321 GTACAGCATCATTTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Db 7703 GCACACCAACACTTCTCATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7644
 Qy 1381 ATGAT 1440
 Db 7643 TTTGAT 7584
 Qy 1441 GAT 1500
 Db 7583 GAT 7524
 Qy 1501 CATTTACCAAGATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 7523 CATTTACCAAGATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7464
 Qy 1561 AGCGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 7463 GTTGCCCAATTTACCAATTTACCAATTTACCAATTTACCAATTTACCAATTTACCAATTTAC 7404
 Qy 1621 CACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 7403 GACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7344
 Qy 1681 TTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1690
 Db 7343 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7334

RESULT 14
 AF064791 1893 bp DNA linear BCT 31-DEC-2001
 LOCUS AF064791
 DEFINITION Pasteurella multocida thymidylate synthase (thya) gene, complete cds.
 ACCESSION AF064791
 VERSION AF064791.1 GI:18000047
 KEYWORDS
 SOURCE
 ORGANISM
 Pasteurella multocida
 Pasteurella multocida
 Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 Pasteurellaceae; Pasteurella.
 REFERENCE
 Authors Moreno J.A., Bosch M., Badiola I., Llagostera M. and Barbe J.
 Title Direct Submission
 Journal Submitted (12-MAY-1998) Departamento de Genetica y Microbiologia
 (Unidad de Microbiologia), Universidad Autonoma de Barcelona,
 Edificio Cn, Bellaterra, Barcelona 08193, Spain
 Location/Qualifiers
 1..1893
 /organism="Pasteurella multocida"
 /mol_type="genomic DNA"

```
gene      /strain="PM25"  
          /specific_host="bovine"  
          /db_xref="taxon:747"  
          /gene="thyA"  
CDS       826..1677  
          /gene="thyA"
```

```

/strain="PM2"
/specific_host="bovine"
/db_xref="taxon:747"
826. .1677
/gene="thya"
826. .1677
/gene="thya"
/EC_number="2.1.1.45"
/codon_start=1
/transl_table=1
/product="thymidylate synthase"
/protein_id="Fai54880.1"
/db_xref="GI:180000948"
/translation="MRYQVLCQRIYDQGWYENPRTGRCITVDADLYDVANNOF
PLIRKSTYKNAALAEFLGSTRBYDNLAAERKQCTGTMMANENSAVLPNRKGD
DMGVTGQGWAKRSGEIVDQRIKTVNNRNGCIDDRKEINTFPNPEFLGRCRD
MHTVTLLELLETLTISYQSCDVPGLNNOIQVFTALMAQTGGKAAKHKIKI
NAHLYEVDLIMKRVQDKRPFPLPOLHINPFLIKLEDELTWYMDRDFVQCHD
IKYFVSQ"

```

ORIGIN

Query Match	26.6%;	Score 775.2;	DB 1;	Length 1693;
Best Local Similarity	67.9%;	Pred. No. 7.8e-229;		
Matches 119;	Conservative	2;	Mismatches 505;	Indels 22;
				Gaps 2;

QY	46	TGACCCCGGATTTGGTTTCGATCGGCCCTTCAGCGGTGGCGGTATGGCTTGTATGTA	105
Db	49	TGATCCCGTCATTTTGGAAATCGGTCTATCGGTTTACGTTGATGGCTTAATGATCT	108
QY	106	GGTGGTTTCCTTTTTGTCTATGTGGTTCGCAATCCCGACGCGATGCGCGGCGATGG	165
Db	109	GCCTGGGGTTTCTCTTTCCTTGGCCCGTTGGCTTCGCGGTGAAAGTGTCCTAACCAACAGCGAGTGG	168
QY	166	TTTGAGCGGTGAGCAAGTCTCGATTGTATATGCGCGGCTTTTATAGTGTGATGATCGG	225
Db	169	CTGGAACAAGACAAAGTGAATAGCTCTGTATTTACGGTTTATGGCGCTTTTTTATAG	228
QY	226	TGGCGAGTTGGTTATGTGATCTTCTACAAATTTGATCTGTCTTGCTGACCCCTCTTA	285
Db	229	TGGCGCATTTGGTTACGTTGTGTTTATTCAAATTCGATTATTTCTTACAAATCCTCCGTA	288
QY	286	TTTATTCAAAGTGGACATCGCGCGGATGCTCCATCCACAGCGCGGCTATTGGGTGTGATCAC	345
Db	289	TTTATTCGCTTTTGGAAAGGCGGATATGCTTTCCATGCGCGCTTAATGTGTGATTTT	348
QY	346	CGGCATGTTCTGGTATGCGCGGTAAACCAACGACCTCTTGGTGGTGGCGGATTTGT	405
Db	349	AGCCATGTTAATACAGGCTTAATAATCAAAACGGGGTTTTCGCAACAGCGGATTTGT	408
QY	406	TGCCCCCTTAATGTCACATTTGGTTTGGGATGGAGATATGGTTAACTTTATGAAATAGTA	465
Db	409	GGACCGTTGATTTCCATTTTGGCTTATGGATGGAGCGGATCGGCAACTTATACATATACGA	468
QY	466	ACCTTGGGAGCAGTAAACGATGTGCTTGGGCTTTGTATTTCCCTAATGTATGGCCACCT	525
Db	469	ATTGTGGGGCGGTGTACCGGATGTCCATGGGCACTGTATTTCCCAACAGCGGATTTT	528
QY	526	GCCGCGCATCCTTACACACTTTATGAAATTTGCGCTTAGAAGCGCTGTCTGTCTTTAT	585
Db	529	ACCACGGCATCCMTTACAGCTATACAGCGGATTTAGAAAGGCATTTGTCATTCCTTAAT	588
QY	586	TCTTAATTTGGTTATTTGTTAAACCTGTCCGCTAGGACGCTATCCGGACTGTTTTAGC	645
Db	589	CTGGAATTTGTATATCAAAAACCTCGACCAATCGGTGCAACGCTTGGTTATTTCTTACT	648
QY	646	TGGATACGGTACATTCGCGTTCCTTGTGAAATACGTCGAGACCAATGCTCAATTGGG	705
Db	649	CGGTTACGGATTTTCCGTTTATCGTTGAGTTTTCGTTAGGCTATGCCAATATAGG	708
QY	706	TCT---CTTTGGCTTCAATTCATATGGGGAAATCCTCCCTACTATCTATGGATGCAT	762
Db	709	CTTATATTTTGGCAACACATCTCTATGGGAAATTTCTTCCACACCAATGATTTCTCAT	768

QY	763	CGGATATTTATGATGATGGTTGGTCTTTCACAGCGGGTTTGTATCAGACCGTGTAGACG	822
Db	769	TGGGCGCGTATCATGTTAGTGGCATATCAAGT-----CGGATA	809
QY	823	AAATATAGGATGTTAGGTGAAACGATATTTAGATCTTGTACAGCGCATTCGTGCATCAAG	882
Db	810	AAAAAAGTAAATTTTATGAAACCTCTTTTAGAATTGTGCCAACGATTCGTATGTAGAG	869
QY	883	TGTTTGGGTTGAAATATAGAGAACGGGCAAGCGTTGTTGACTGTGATTAATGCGCATTT	942
Db	870	TGTTTGGGTTGAAATATAGAGACTGAAAAAGTTGTTTACCGGTATATGACGATCT	929
QY	943	GACCTACGATGTGGGCAACATCAGTTTCTCTTAGTACTTACACGCAAGCTTTTGGAA	1002
Db	930	CACCTACGATGTGGGCAACATCAGTTTCTCTTAGTACTTACACGCAAGCTTTTGGAA	989
QY	1003	AGCTGCCGTAGCCGACGTTGCTCGGCTATTTTCGGTGAAGTATAGATATGCGGCGATTTTCG	1062
Db	990	AGCGGCATGTGCAGAAATTTTATGATATATTCCTGGCTATAGACATATGCGGCAATTTCCG	1049
QY	1063	CCAAATTAGTATCCAAAACCTGGATGTCTAATGCCAATTTAAACAGCATGCTCTCAAA	1122
Db	1050	CAAACTGGGACCAAAAACCTGGGATGGCAATGCTAATGAAAAAGTGCATGGCTCAATAA	1109
QY	1123	TCCTTACCGTAAAGTGAAGATGACATGGAACGGGTATAGTGTTTGAAGGTATAGCTTG	1182
Db	1110	TCCGACACCGCAAGACCGCATGATATGGAACGTTTATAGCGCTACAGGCGCTGCATG	1169
QY	1183	GGCTAAGCGTATGTTGGTCAATTTTGAACCATTTGAAAAAAGTTGTATGATTTTGAACCG	1242
Db	1170	GCGCAACCGACGCTGAACCCGTCATATTCGCAAAAATGCTCAATATCTGCCAA	1229
QY	1243	TGGCGTGAATGACCGAGGTGAATCTTTPAATCTTACAACTCGGGTGAATTTTCAATGG	1302
Db	1230	CGGATATGATGACCGTGGTGAATCATGACCTTCTTCAACCCCGTGAATTTTGAACCTGG	1289
QY	1303	CTGTTTGGGCTTGCAATGTAACAGCCATCTTTTCAATGGTGGGGATACCTGTACT	1362
Db	1290	TTCCTTAGCTTGTGACATGACACCCACCACTTCTCATCTACTGCTGATACCTCTATCT	1349
QY	1363	CAACAGTACTAGCGCTTCATGTGATGTGCCCTTGGGGTTGAATTTCAACATGTGTGACGT	1422
Db	1350	CACAGTATATCAGGGTCTTGTGATGTGCGCTGGCTTGAAATTTCAACAAATTCAGT	1409
QY	1423	TTATGTGTTCTTTCGCTGATGAGACAGATCACAGGAAAAAGCGGGCTTGGGGTATCA	1482
Db	1410	GTTCACCTTTCTTCTCTGATGAGCAAAATTAACGGGCAAAAAAGCGGGTAAAGGGTACCA	1469
QY	1483	CAAGATCGTCATGCGACATTTTACCAAGATCAACTGCAATTTGATGCGCATGTGCACT	1542
Db	1470	CAAAATATCATGACACACATTTTATGAAGATCAACTTGATCTCATGAAAAATGTGCATTT	1529
QY	1543	AAAAAGTAGCCATTTCCACGCGCTCAGTTTCATATCATATCCAAAGATTTAAAAACATGCA	1602
Db	1530	AAAAAGTAGCCATTTCCACGCGCTCAGTTTCATATCATATCCAAAGATTTAAAAACCTTCA	1589
QY	1603	GGATTTGGAAACTTGGGTCACTTTGATGATTTTGAAGTCAACGGATATCATGTTTCCACGA	1662
Db	1590	AGATTTGAAACTTGGGTCACTTATGATGATTTTAAAGTCAACGGGTATCATATGTCACGA	1649
QY	1663	TCTATATCAATACCGTTTTCAGTCTAA	1690
Db	1650	TGCCATTAATATCTTTCTCTGTGTA	1677
RESULT 15			
LOCUS	U32772	11545 bp	DNA
DEFINITION	Haemophilus influenzae Rd section 87 of 163 of the complete genome.		
ACCESSION	U32772	142023	
VERSION	U32772.1	GI:1573918	
KEYWORDS	Haemophilus influenzae Rd	KW20	
SOURCE			

ORGANISM	Haemophilus influenzae Rd KW20
REFERENCE	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.
AUTHORS	1 (bases 1 to 11545) Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shiley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhmann, J.L., Geohagan, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.
TITLE	Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
JOURNAL	Science 269 (5223), 496-512 (1995)
MEDLINE	95350630
PubMed	7542800
REFERENCE	2 (bases 1 to 11545) Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Bordovsky, M., Rudd, K.E. and Koonin, E.V.
AUTHORS	Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli
TITLE	Curr. Biol. 6 (3), 279-291 (1996)
JOURNAL	96398784
MEDLINE	8805245
REFERENCE	3 (bases 1 to 11545) White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
AUTHORS	Direct Submission
TITLE	Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA
REFERENCE	4 (bases 1 to 11545) White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
AUTHORS	Direct Submission
TITLE	Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA
REMARK	The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
REFERENCE	5 (bases 1 to 11545) White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.
AUTHORS	Direct Submission
TITLE	Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA
REMARK	The whole genome was shifted by 568 nucleotides for a new start
COMMENT	On Sep 30, 1996 this sequence version replaced gi:1221622.
FEATURES	Location/Qualifiers
source	1..11545
gene	/organism="Haemophilus influenzae Rd KW20"
gene	/mol_type="genomic DNA"
gene	/db_xref="taxon:71421"
gene	/complement(122..604)
gene	/gene="HI0899"
gene	/complement(122..604)
gene	/gene="HI0899"
gene	/note="similar to GB:D10483 SP:P00379 GB:J01609 GB:V00276 GB:X05108 percent identity: 53.16; identified by sequence similarity, putative"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="dihydrofolate reductase (folA)"
gene	/protein_id="AAC22559.1"
gene	/db_xref="GI:1573919"
gene	/translation="MTFSLIVATLTNNVIGKDNOMPHLPADLAMPFRONTGKGVING RTFESIGRPLEKRTNIVLSROLPEHGVIGWDSFSAVNFVDFEDTMLIGSGEIKK QYLPKADKLYLTQITBLDGTFFQLMWEWELEFDEYRKADENRDKCFLLITRK "
gene	705..1811
gene	/gene="HI0900"
gene	/note="similar to GB:U00096 SP:P07005 PID:1208988 PID:1552810 PID:1786437 percent identity: 65.57; identified by sequence similarity, putative"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="glutamate 5-kinase (gamma-glutamyl kinase) (proB)"
gene	/protein_id="AAC22560.1"
gene	/db_xref="GI:1573920"
gene	/translation="NNKKTVVFGSTLTQSGPKLNSPPHMEIVRQIAQHNDFRI VITSGIAGRGVYINHPOLPRTIASKGLAAGSGLQAMKRLAIDYHIGOLL TRADIERERFARDTLYLLNHIIPVYNEDAVATAIRKYGDDNNSALVATLYQ AEOYLTLTDOQGLPDSDPKRNPAKULPVEQITDHRSTRISAGSGTNLTGGMATKIT AADVATRSGLIETIAPGNRPNVIAADAYEONIGTKFIHQSDRLSESKQWLFPAPISAG IITIDNGAQNVAILEQNKSLPAGIINVEGFRSGEVVKIRIQSGKDIALGMPRYNSDA LQIKGRKSADINVLGYEYGAVAMERDDMIILS"
gene	1892..2482
gene	/gene="HI0901"
gene	/note="similar to PID:882723 GB:U00096 SP:Q46930 PID:1789194 percent identity: 77.71; identified by sequence similarity, putative"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="conserved hypothetical protein"
gene	/protein_id="AAC22561.1"
gene	/db_xref="GI:1573921"
gene	/translation="MIDPDGVRPNVGIVICNRKQVLMWAKRCQNSMOPQSGINDNE SAEQAVRELEHEVGKOPKDVRLLYSKWMLRPLRLRYSKMCGKQKRWEL LIRSDKNINMOTKSPBPDGMRWSEFMVPRVYVSPKDYRWKMKPEASLIFLTNP LIRSASREANSLHRSANKYSQTKYTKRHFYKSRGQ"
gene	2482..3276
gene	/gene="HI0902"
gene	/note="similar to GB:L77117 SP:Q57883 PID:1591145 percent identity: 28.84; identified by sequence similarity, putative"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="conserved hypothetical protein"
gene	/protein_id="AAC22562.1"
gene	/db_xref="GI:1573922"
gene	/translation="MFTFILLCLVGLAGLAFGLGGGLVPTVLYLPIYVNP ESLLMSTALSTSPATVITVIGSGAQHHKGNIVQAVYILAVYLSFTICGFLTGR LDEIISAKTIPACLVVYLATRWVLSIRKDDVTKSLPPLSSVIGIILGNASSAAGTGG GFIIVFLPARGINIKQAIGSSAFCGMLGISGMPFIYSGNGNPLMPSPISGLYILP AVLGITATSFETSFKGASATAPKLPVTLKGFALFIVVAIINMFLK"
gene	3285..4091
gene	/gene="HI0904"
gene	/note="similar to SP:P37149 GB:U12289 PID:516036 PID:882720 GB:U00096 percent identity: 62.78; identified by sequence similarity, putative"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="prolinoprotein diacylglycerol transferase (lgt)"
gene	/protein_id="AAC22563.1"
gene	/db_xref="GI:1573923"
gene	/translation="NNSNYLLLPFPDPSIFTLGDSNIGRWGLNYLLGFVPAWLVAV RANRPNSGATVDVDSLFGNGMGVFIAGVGVDFVYLDHQLQPLVLFVWEGGM SFHGGLIGVIVAMITVTSYQKXNFQVOTAFVLPFGGDLGILNINLEIEMGRGTN VPVAMTFPNDPLLPHPSPQLYEAPLEGLVFLTINIPIKKRPAPASVAGFLIGYGV FRVIVVBPBPVENFPGIITGQALCLPMIIGATIMMAWAKSAVVK"
gene	4101..4952
gene	/gene="HI0905"
gene	/note="similar to GB:J01710 SP:P00470 PID:147987 PID:42685

PID:882719 percent identity: 35.32; identified by sequence similarity; putative"

/transl_table=1
/codon_start=1
/product="thymidylate synthetase (thyA)"
/protein_id="AAC2564.1"
/db_xref="GI:1573924"

/translation="MKQYLRICRRIYVSEGWANPERGKGLTYINADLEVDVNNQF
PLITRKSYWKALAEFLGIRGIDNADFPALGKTKTDANANEMANLPHRGSD
DMGRVYGVQGRAMKRENGEITDQLEKTIYNVNLKGI DDEGALMITEFNGEPLGLRGC
NMHTTFSLVGDTLHLSYQSCSDVPLGANTFNOIQVFFFLAMQITKSKGAKVHKV
NAHIVEDQLEIMRDVQLKREBPPLPKLEINPDKITLEDLETFWVDDFKVGYOSHEP
IKYPSV"

4952 5473
/gene="H10906"
/note="similar to SP:P30134 PID:296183 PID:387637
PID:1033148 GB:U00096 percent identity: 57.41; identified
by sequence similarity; putative"

/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAC2565.1"
/db_xref="GI:1573925"

/translation="MDAAKVRSEEDKMMRYALIELADKRAALGEIPVGAIVDDARNI
IGEMNLSIVOSDPTAHAEITLRGAKNIQVRLNLTLYLEPCTMGALIHRS
IKRIYVGSADYKTAIGSRFHFDDYKNNHTLEVTGVLAECSGKSTFQKRSEK
KIEKALIKSLSDK"

complement(5501, 5815)
/gene="H10907"
/note="hypothetical protein; identified by Genemark;
putative"

/codon_start=1
/transl_table=1
/product="H. influenzae predicted coding region H10907"
/protein_id="AAC2569.1"
/db_xref="GI:1573930"

/translation="MENKAERYOKAVNITDVLQSPFAKIIKKGLAINEINQENRIF
POEFHGRIGNMTDINSIFIEPANAIVRGILFRQTELLKLIQEEFPQVTFEITLNP
GR"

5895. 6197
/gene="H10908"
5895. 6197
/gene="H10908"
/note="hypothetical protein; identified by Genemark;

Query Match 25.6%; Score 744.2; DB 1; Length 11545;
Best Local Similarity 66.7%; Pred. No. 4,9e-219;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

QY 2 AGAAGGTTGTTATGCTCAGGTTATCTGAGTTTCCCATATTTGACCCCGTATTTGTT 61
DB 3273 ATAAAGAAATTTATGATTCATTAATTTACTTCTCCCACTTGATCGAGTATTTT 3332
QY 62 TCGATGGC-----CCTCTAGCGGTGCGGTGATGCTTATGATGATTTGGTGGTTTC 115
DB 3333 ACCCTTGGCGATAGTAATATCGGCTTACGTTGATGAGCTTATGATGATACCTTTAGGTTT 3392
QY 116 CTTTGGCATGATGTTGGCAATCGCCGAGCGATCGCGCGGAGAGTGTGACGCGT 175
DB 3393 GTTTTGGCAGTGGCTGGCTGGCTGGCTGCTAAACGCCCAAAAGACGTTGGACAGTA 3452
QY 176 GAGCAAGTCTGACTGTTGATTCGCCGCTTTTAAAGTGTAGTATCGTGGCGAGTT 235
DB 3453 GATCAAGTGTATGCTTACTTTCAACGGTTTATGGGGGTGTATTATGGCGAGCGTGT 3512
QY 236 GGTATGATCTCTACAAATTTGATCTGTTCTGCTGACCGCTTATTTATTTAA 295
DB 3513 GCGGATGATTTTCTATATCTGATCATTTCTTCAAGAAACACTTTATTTATTCGCG 3572

QY 296 GTGTGACTGGCGGATGTCCTTTCACGCGGCTTATTTGGTGTGATCAACGCCATGTTTC 355
DB 3573 GTTTGGAAGTGGAGATGTCCTTTCACGCGGCTTATTTGGTGTGATGTCATGAT 3632
QY 356 TGGTATCGCGTAAACCAACGACCTTCTTGTGTGCGCGGATTTTGTCCCTTTA 415
DB 3633 TGGACATCTTATTCGAAAAGCTATTTTGGCAACGCGGATTTTGTGCGCTTTG 3692
QY 416 GTGCCATTTGGGTTGGGAGTGGAGCGTATCGGTAACCTTATGAAATGMACTTTGGGA 475
DB 3693 ATTCCGTTGGTTTGGTTTGGGCAAAATGGTAAATTCATTAATCTTGAATCTTAAGGGA 3752
QY 476 CGAGTAACGAGATGCTTGGGCTTTGATTTCCCTAATGTGGCC--CACTGCCGCC 532
DB 3753 CGGAAACGAATGCTTGGGCAATGATTTTCCGAATGATCCTTTTACGCTCGT 3812
QY 533 CATCTTCACGCTTATGAAATTCGCTTGAAGGCGTCTGTCTTTATCTTAT 592
DB 3813 CATCTTCACGCTTATGAAATTCGCTTGAAGGCGTCTGTCTTTATGAAATTCGAT 3872
QY 593 TGGTTATTTGTAAACCTGTCGCTAGGACGATTCGGAATGTTTAACTGATG 652
DB 3873 ATTTTATTAATAAACACGTCATGCTTGTGACGTTTATCTTAATGCTTAT 3932
QY 653 GGTACATTCGCTTCTTGTGAAATCGTCGAGGCGGATGCTCAGTTGGTGT 712
DB 3933 GGGGTCTCGTTTATTTGTGAATATGCGGTAACCTGAAG-----TTGAAATTTTC 3986
QY 713 GTGGCTCATTTCAATGGGCGCAATCTCTCTTAACCTATGATGATATGATTTTG 772
DB 3987 TTTGGATATATACAGAGGCGAGCCCTTGTGTCGATGATATGTTGGTGTCTTC 4046
QY 773 ATGATGTTTGTCTTACAGAGCGGCTTGTATCAAGACCGTGTAGACCAAAATAGGT 832
DB 4047 ATTTATGCTTGGCTTATTCACGC-----AAAAGTGGCTATTAATAATGGAG 4094
QY 833 AGTTAGTGAAACAGATTTAGATCTTGTACAGCGCATGTCGATCAAGTGTGGGT 892
DB 4095 ATTTTATGAGCAATATCTGAGCTTGTGCGCGATTTGATAGTAAAGGAGTGGT 4154
QY 893 GAAATGACGAACGCGGCAAGCTTGTGTGATCTGTATTAATGCCATTTGACCTACGT 952
DB 4155 GCTAATGACGTAACAGTAAGCATTCCTCACGCTGATTAATGACATTTAGAAATGAT 4214
QY 953 GTGGGCAACATCAGTTTCTCTAGTACATCAACGAGATTTTGAAGTGGCGTA 1012
DB 4215 GTGGCAATATCAATTTCCGCTGATTACTACCCGTAAGATTTTGAAGACGCGATT 4274
QY 1013 GCCGATTTGCTCGCTATATTTGTGTATCGATATATGCGCGGATTTGCCATTTAGT 1072
DB 4275 GCTGAATTTTATGTTATATGCTGATATGACATGCGGATTTCCGCGCACTTGGC 4334
QY 1073 ACCAAACCTGGATGCTAATGCCATTTAAACCAAGCATGCTCAACATCCTTACGCT 1132
DB 4335 ACGAAACCTGGATGCTAATGCGAATGAAATGACGCTTGGCTTGAATTCGATCGT 4394
QY 1133 AAAGTGAAGATGACATGAGACGCGTATGTTCAAGGTATGACCTTGGCTTAAGCT 1192
DB 4395 AGAGCGCTGATGATATGAGGCGCGTATATGATGTCAGGACGACATGCGTAAGCT 4454
QY 1193 GATGTGTGATATGACCACTTGAAGAAAGATTTGATGATTTGACCGCGCTGAT 1252
DB 4455 AATGGAAGAACTATGATGATGCTACGTAAGTAATTTGTAATACTTAAGAAAGTATGAT 4514
QY 1253 GACCGAGGTGAATTTCTTCAATCCGCGTGAATTTCACTGCGGTGTTGGCG 1312
DB 4515 GATAGAGGAAGATTTTAACTTTTAAACCTCGGGGAATTTGATCTTGGTGTCTTCGT 4574
QY 1313 CTTTGATGTACAGCAATTTTTCATTTGCTGGGAGTACCTTGTATCTCAACACTACT 1372
DB 4575 CTTGTATGATGACGATCTTTTCTCTGTGGGAGTACTTTACATCTTACTACTAT 4634
QY 1373 CAGGTTATGATATGAGCCCTTGGGTTGAATTTCAACATGTTGAGGTTATGTTCT 1432

```

Db 4635 CAGCGTTCTGTGATGTTCCCGCTTGATTAATTCATCAATTCAGGTGTTACCTTC 4694
QY 1433 CTGGCGTGTATGGACAGATCACAGGAAAAAGCCGGGCTTGCGTATCACAGATCGTC 1492
Db 4695 TTAGCACTTATGGCTCAGATCACGGGCAAAAAGCCGGGAGCATATCATTAATTTGTG 4754
QY 1493 AATGCGCAATTACCAAGATCACTCGAATTGATGCGGATGTGCAAGCTAAACGTGAG 1552
Db 4755 AATGCGCAATTATTAAGATCAGCTTGATGATGCGTACGTAACAATTAAACGCGAG 4814
QY 1553 CCATCCCGCAGCGCTCAGTTCATATCAATCCAAAGATTAAACACTGAGGATTTGGA 1612
Db 4815 CTTTCCCTTTACCAAAATTAGAAATTATCCAGATATMAAACGCTTGAGATTTGGA 4874
QY 1613 ACTTGGGTCACTTTGGATGATTTTGAAGTACCGGATATCAAGTTCACAGATCTATTCAA 1672
Db 4875 ACTTGGGTCAAGATGATTTTAAAGTGTGCTATCAATCCACGAACCAATTAA 4934
QY 1673 TACCGCTTTCAGCTAAT 1691
Db 4935 TATCCTTTTCCGCTAAT 4953

```

Search completed: July 31, 2004, 15:38:59
 Job time : 7462 secs

